

RESULT 13
AF384033 linear PLN 26-JUN-2001
LOCUS 1124 bp mRNA
DEFINITION Zea mays HD2 type histone deacetylase HDA106 (hda106) mRNA,
complete cds.
ACCESSION AF384033
VERSION AF384033.1 GI:14550105
KEYWORDS Zea mays.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1124)
AUTHORS Chandler,V.L., Kaeppeler,S.M., Kaeppeler,H.F. and Cone,K.C.
TITLE Sequences from the Plant Chromatin Consortium (NSF Plant Genome
project 9975930)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1124)
AUTHORS Bergstrom,D., Springer,N.M., Schmitt,L.T., Guthrie,E.,
Sidorenko,L., Selinger,D., Kaeppeler,S.M. and Cone,K.C.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-2001) Biological Sciences, Univ. of Missouri, 101
Tucker Hall, Columbia, MO 65203, USA
FEATURES Location/Qualifiers
1..1124
 `/organism="Zea mays"`
 `/cultivar="B73"`
 `/db_xref="taxon:4577"`
1..1124
 `/gene="hda106"`
120..980
 `/gene="hda106"`
 `/codon_start=1`
 `/product="HD2 type histone deacetylase HDA106"`
 `/protein_id="AAK67143.1"`
 `/db_xref="GI:14550106"`
 `/translation="MRVGGQEVKPGATVSCVKVDGLVIHLSONALCESKKASENAILS
VNIDKKLVIGLTSVEKHPIISCDLVDFDKDPLPHNSKTRSVFFRGYKSPVPFLFSNS
GESSDESLKTDOIPLONNEIKISAAYPAKDODDDVFILAMMMIYSDDDDDDFT
TSDSNSEESDDSDDESDDSGADPDSDDSGESTSAPKRTD
VVVGKRRAIKAEAPYGKKAKSEQSQKGTGDKASTSHPAKOSIKTPADKSRKTKPADK`
BASE COUNT 327 a 237 c 265 g 295 t
ORIGIN
Query Match 11.6%; Score 109.2; DB 8; Length 1124;
Best Local Similarity 55.2%; Pred. No. 5.8e-14;
Matches 235; Conservative 0; Mismatches 188; Indels 3; Gaps 1;
OY 41 ttctagccatggagtctgtggggaattgaagtttaaatcaggaaagccaggtcacagtgc 100
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 TCGGCCAATGAGGTGCGGGGTCAAAGAGTAGAAGCCAGGACCGGTTCTTGCAAAG 171
OY 101 ctgaagaaggcatcttatccagtttctcagcgatcgttgtagagaatgtataaacacaaga 160
|| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 TTGGTCATGGCTTGGTATTCCACCTTTACAGAGCTGCAC'TGGGGGAATC---AAAAGAAG 228
OY 161 agggagattgtgcccttacatgtaaaagttgggaaccagaactggttcttggaactc 220
|| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 229 CGAGTCAGAA'GCCATTTC'TGCTGTAATAATCATGATAAGAAC'TAGTCTTGGAACCT 288
OY 221 tatcgactgagaacatccctcaagctttcttgtgattgttcgcagaaaggaggtttgagc 280
|| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 TATCAGTTGAGAAGCATCCTCAGATCTCTGTGATCTTGATTGATTAAGGATTTTCAGC 348
OY 281 ttctcacacttgggaaaaaggaagtgttacctttgttgatatacaaaaactccccacattg 340
|| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 349 TACCACAACAA'TTCAAAACACG'TAGCGTTTTCTTCGCTGGCTACAAGTCGCCCTGTCCTC 408
OY 341 agccacaaggctattctlgaggagaagagagagagagagagagagagagagagagagagag 400

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Eumariophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza. 1 (bases 1 to 894) Dangl,M., Brosch,G., Haas,H., Loidl,P., Graessle,S. and Lusser,A. Molecular characterization of type-2 histone deacetylases in higher plants JOURNAL Unpublished REFERENCE 2 (bases 1 to 894) Dangl,M., Haas,H., Loidl,P. and Lusser,A. Direct Submission JOURNAL Submitted (13-APR-2000) Microbiology, University of Innsbruck, Fritz-Pregl-Str. 3, Innsbruck 6020, Austria
FEATURES source location/Qualifiers 1..894 /organism="Oryza sativa" /db_xref="dbEST:D15380" /db_xref="taxon:4530" 1..894 /codon_start=1 /product="putative histone deacetylase HD2" /protein_id="AAF70196.1" /db_xref="GI:7804960"
CDS 1..894 /translation="MEFWGLVKGPGTVCPEDEREFLHLSQAALGESKSGSDNAVHY VKTDQKLVIETLSADKPQIOFLDFKEFLSHTSKTASVFFSGIKYVSQPAEDEHM FDSSVEDEEEEEKIILPAPRAKVGEKENEQKOGTKDTSSAKSAVNDDDDDDDD SAEDSDQEDLSPPDDDDSSDEDDEDEPTPKPETCKRKVAIEVLKTPS SDKKFIATPSQKTKDKKKGVHVATPHPAKOAKSTPVNDKSKSEKSPKSGGGSISCKSC SKTFNSEMALQSLSKAKHPAK"
BASE COUNT 289 a 172 c 234 g 199 t
ORIGIN
Query Match 12.5%; Score 117.6; DB 8; Length 894;
Best Local Similarity 55.3%; Pred. No. 8.4e-16;
Matches 272; Conservative 0; Mismatches 214; Indels 6; Gaps 2;
QY 49 atggaattctgtgggaattgaagttaataatcaggaaagccagtttacatgactcttcgaaga 108 Db 1 ATGGAGCTTCGGGTCTTGAAAGTCAAGCCTGGACAGACTGTCAAATGTGAGCCCTGAAGAT 60 QY 109 ggcattcttaccagttctcagcgcatcgcttggaagtgtataaacagaagggagag 168 Db 61 GAACGCCTTTTGCACCCTTCTCAGCGTGTCTTGGGGAAAT--CAAGAAGAGATCTGAC 117 QY 169 ttgtgctttacatgtaaagttgggaaccagaacttgttctgggaactctatgcagt 228 Db 118 AATGCAGTAATGATGTATAAAGTAGATGATCAAAAGTAGTCAATCGAACCCCTCAGCT 177 QY 229 gaaaacatcccctcagcttttctgtgatttggtattgcacaagagatttgagctttctcac 288 Db 178 GACAAGTTCCTCCAATCCAGTTTGATTGGTCTTTGACAAAGAGTTTGAGCTGTGCACAC 237 QY 289 acttggggaaaaaggaagtgttacttgttgatatacaaaactccccacatttgagccacaa 348 Db 238 ACTTCAAGAGCTGCTAGTGTGTCTTTTCTGGCTACAAAGTTTCCCAGCCGCTGAGGAA 297 QY 349 ggctattctgaggaagaagagggaaggaaggaagtagttctctgtgggaatgctgccaaag 408 Db 298 GATGAATGGATTTTGATTCTGAAGAAGTTGAAGATGAAGAGGAGGAAGAAAAGATCATT 357 QY 409 gctgtagctaaaccaaggctaagcctgcagaagtgaagccaagcttggatgatgaagag 468 Db 358 CCAGCTCCCGACGGCCAAATGGCAA---AGTTGAAGGGAAGGAAGAAATCAGCAGAAAAACAA 414 QY 469 gatgaagctgattctgacggaattggatgaagatgattctgatggttgaggattctgggaa 528 Db 415 GGCAACACAGATTCTTCAGCTTCAAATCAAAGCGTCGAGTGAATGACNGATGATGATGAT 474 QY 529 gaagagcctaca 540 Db 475 GATGACAGTGTCA 486

FEATURES		Location/Qualifiers	
source		1. .885	
		/organism="Arabidopsis thaliana"	
		/db_xref="taxon:3702"	
CDS		1. .885	
		/codon_start=1	
		/product="putative histone deacetylase HD2c"	
		/protein_id="AAF70197.1"	
		/db_xref="GI:7804963"	
		/translation="MEPWGVKNGKPLHLDPGLDRLVHISOVALGESKNVTEPIQLVTVGSKLLIGTILSHEXFPOLSTFIVLERNFALSHTKMGSVFTSGYKVDAADPEPEDLIDDLEAAGFKAPKSAAKVNFOLNEDVKAQODDADGSEDSDDDDSENGSGDEEEKVTAESSEDDSDDEDDSSEREETPKPEPKRSAPKSNKPSANPKAKAFVTPQMDKKPHVHVAITPSPKQAGKNSGGSGTGETSKQQQTPKRSAGAGFGCKSCRTTF	
BASE COUNT	266 a	190 c	231 g
ORIGIN		198 t	
Query Match	13.1%	Score 123.4;	DB 8; Length 885;
Best Local Similarity	51.2%	Pred. No. 4.4e-17;	
Matches 289;	Conservative	0; Mismatches 276;	Indels 0; Gaps 0;
QY	49	atgagctctcgggaattgaagttaactcagaaagccagttacagtgactctcctgaagaa	108
DB	1	ATGGAGTTCTGGGCTGTTGAAGTTAGATGGCAACCACTTCACCTAGCTCTCGGCTA	60
QY	109	ggcattcttaccacgtttctcaggcatcgcttgggaagtgtataaacagaagggagag	168
DB	61	GACAGGCTCGTGACATCTCTCAGGTGCTTGGGAGAGATGAAGAACATGTAAACCGAA	120
QY	169	tttgtaccttacctgaagtgtgggaaccagaacttggctcgggaactctatcgact	228
DB	121	CCAATCCAGCTTTACCTGACGTGTTGATCGGATAAGCTTCTCAATTGGAAGCTATCTCAT	180
QY	229	gaagacatccctcagctttctgtgattggtattcgacaagagtttgaqctttctcac	288
DB	181	GAGAAGTTTCTCGCTGTCTACGGAGATGTCCTTGGAAAGGAACCTTTGGCTGTCTCAT	240
QY	289	acttggggaaaaaggaagtgtttacttttgggtatatacaaaactccccacttgggccaaa	348
DB	241	ACTTGGAAAGTAATGGGAGCGCTTTCTCTCTGTTGTACAAAGTTGATGATCGCGAG	300
QY	349	ggctattctgaggaacaaaggaaggaaggaaggaaggttcctcgtcgggaatgctgccaag	408
DB	301	CCTGAGATTTGATTGATGATCAACTTTGAGGCTGCTGGTTTCAAAGCTGCTCGGAAATCT	360
QY	409	gctgtagctaaacaaaaggcctaagcctgcagaaagtcagaagtcagcagctgtgatgatgaagag	468
DB	361	GCTGCGAACAAGCTGAACCTTTCAGTTGCCAATGAAGATGTCAAAGCCAAGCAAGATGAT	420
QY	469	gatgagctgattctagcgaagtggatgaagatgattctgatggtgaggaattctcgaggaa	528
DB	421	GACGCTGACGGTATGTAAGAAGAAATCTTTCAGATGATGATGATAGTGAAGAACTCTGGAGAC	480
QY	529	gaagagcctcacactaagaagcctgcatacgaacgaagaagagactaatgaaactaccct	588
DB	481	CAGCAGGAAGAGAAGGTTACTGCTGAATCTGACAGTGCAGGAAGATGACTCATCAGACGAC	540
QY	589	aaagcacctgtgtcagcaaaagaag	613
DB	541	GAAGAAGATGACTCCTCAGAAAGAGG	565
RESULT	12		
AF255711		894 bp	mrna linear
LOCUS			PLN 15-MAY-2000
DEFINITION			Oryza sativa putative histone deacetylase HD2 mRNA, complete cds.
ACCESSION			AF255711
VERSION			AF255711.1
KEYWORDS			GI:7804959
SOURCE			Oryza sativa.

LOCUS	AF044914	918 bp	mrna	linear	PLN 11-FEB-1999			
DEFINITION	Arabidopsis thaliana putative histone deacetylase (HD2) mRNA, complete cds.							
ACCESSION	AF044914							
VERSION	AF044914.1	GI:2854069						
KEYWORDS	thale cress.							
SOURCE	Arabidopsis thaliana							
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 918)							
REFERENCE	Dangl, M., Haas, H. and Loidl, P.							
AUTHORS	Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2							
TITLE	Unpublished							
JOURNAL	2 (bases 1 to 918)							
REFERENCE	Dangl, M., Haas, H. and Loidl, P.							
AUTHORS	Direct Submission							
TITLE	Submitted (27-JAN-1998) Microbiology, Medical School, University of Innsbruck, Fritz-Pregl-Str. 3, Innsbruck 6020, Austria							
JOURNAL	Location/Qualifiers							
FEATURES	1..918							
source	/organism="Arabidopsis thaliana"							
	/cultivar="Columbia"							
	/db_xref="taxon:3702"							
	/clone="133F157"							
	/clone.lib="EST-clone 133F15 (T. Newman, Plant Physiol. 106, 1241-1255, 1994)"							
gene	1..918							
	/gene="HD2"							
CDS	1..918							
	/gene="HD2"							
	/codon_start=1							
	/product="putative histone deacetylase"							
	/protein_id="AAC02539.1"							
	/db_xref="GI:2854070"							
	/translation="MEFGWAVTPKNAKVTPEDSLIVHISOASLDCTVKGSGSVLSVTYQGLKLVIGTSLQSPDLVDFKFEFSLHSGTRKANVHFICGVKSNIDDDFTSSDDEVPDPAVPATVATVANGAAGVVKADTKPAKPAEVKPAEKPESDEEDSDDEDEEDDSEKGDWDDSDDEEDSEDEBEETPKPEPINKKRPNESVSKTVSGKAKAPAAATPSTPQTEKKKGHTATPHPAKKGGKSPVNAOSPKSGCGGSSGN							
	NNKKPPNSGKFGGSGNNKSGNKGGKRA"							
BASE COUNT	292 a	194 c	239 g	193 t				
ORIGIN								
Query Match	16.08; Score 149.8; DB 8; Length 918;							
Best Local Similarity	64.8; Pred. No. 6.6e-23;							
Matches 239; Conservative	0; Mismatches 127; Indels 3; Gaps 1;							
Qy	49 atgagttcttgggaattgaagttaaatcaggaaagccaggtacagttccttgaagag 108							
Db	1 ATGAGTCTCGGGAGTTCGGGTGACACCAAAACCGTACTAAGGTGACTCTCGAAGAA 60							
Qy	109 ggcattctatccacgtttctcaggtatcgcttgagaagtataaaacaagaggagag 168							
Db	61 GACAGCTTGTCCACATTTCTCAGGCTTCACTT--GACTGCACAGTGAATCTGGAGAA 117							
Qy	169 ttctgctttacatgtaaagggttggaaacagaaactgttctgggaactctatcgact 228							
Db	118 TCTCTGCTTTTGTGTCGACGTGTTGGTGGGCTAAACTTGTATTGGAACACTTTCACAA 177							
Qy	229 gagaacatccctcagctttcttgatttgatttcgacaagaggatttgactttctcac 288							
Db	178 GACAAGTTCCTCCATAGTACGTTTCATTTGGTTTTTGTATAAAGAGTTTGACCTTTCACAC 237							
Qy	289 acttggggaaaaggaagtgtttactttgttgatcacaaaactcccaactgagccacaa 348							

Db 298 CACTTCACTAGTGGGATGATGAGGATGTTCTCTGAAGCTGTCTCTGCTCTGCCCCCTACT 357
 Qy 409 qctqtagct 417
 ||||| ||
 Db 358 GCTGTTACT 366

RESULT 9
 AF372889
 LOCUS Arabidopsis thaliana AT5g03740/F17c15_160 mRNA linear PLN 23-MAY-2001
 DEFINITION Arabidopsis thaliana AT5g03740/F17c15_160 mRNA, complete cds.
 ACCESSION AF372889
 VERSION AF372889.1 GI:13926267
 KEYWORDS FLU CDNA.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1236)
 Authors Chew, R., Chen, H., Kim, C.J., Meyers, M.C., Shinn, P., Banh, J.,
 Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 TITLE Arabidopsis cDNA clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1236)
 Authors Chew, R., Chen, H., Kim, C.J., Meyers, M.C., Shinn, P., Banh, J.,
 Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (18-APR-2001) Salk Institute Genomic Analysis Laboratory
 (SIGAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the
 sequencing and annotation of the RAPL cDNAs: Shinn, P., Chen, H.,
 Chew, R., Kim, C.J., Koeseema, E., Meyers, M.C., Tracy, S.E., Banh, J.,
 Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
 Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
 and Ecker, J.R.
 Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 FEATURES
 source
 1. .1236
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAPL03-05-D06(R12384)"
 /note="ecotype: Columbia"
 1. .112
 113. .997
 /note="histone deacetylase-like protein"
 /codon_start=1
 /product="AT5g03740/F17c15_160"
 /protein_id="AAK49605.1"

5'UTR
CDS

Db 61414 TTGTTATAGCTTCTCTGATTTGTCCTCTTTTGTGTTACCATTTTCAGATAAATGACTGAGG 61355
 QY 490 ----- 489
 Db 61354 AGATGAGACATCTTTTTCACAAATGTTGGTGACGATCTTCTGTATGATCTTGCA 61295
 QY 490 -atggaatgaagatgattctgagtgagatctctgaggaagagcctacacctaagaa 548
 Db 61294 GATGGATGAAGATGATTTCTCATGCTGAGGATTTCTGAGGAAGAGCCCTACACCTAAGAA 61235
 QY 549 gctcgcatacaagaagagcttaataaactccccctaaagcaccttgtatcaagaaa 608
 Db 61234 GCTGATCATCAGCAAGAGAGCTNATGAACTACCCCTAAGCACCCTGTCTCAGCAA 61175
 QY 609 gaaggcgaagtgcagttactcctcagaaa ----- 639
 Db 61174 GAAGCGAAAGTAGCAGTTACTCTCTCAGAAACAGGTATAGATACAAACCATCTCTTAC 61115
 QY 640 -----a 640
 Db 61114 ATTATTAATAATGGGAACACAGGATTCATTTGAACCTGCATCGTCTTTTGGCA 61055
 QY 641 cagatgagaagaaggggaaagcgtgcaaacacagagcccaagtcgagccagtcag 700
 Db 61054 CAGATGAGAGAGAAAGGGGAGGCTGCACAGAGCCCAAGTGGCCAGTCAAG 60995
 QY 701 tctcagtggttctcatgcaagaa ----- 722
 Db 60994 TCTCATGTGTTTCATCAAGAGTGAGACTCTCTCTTTCTTAACCTCTGTCTTCACT 60935
 QY 723 -----gac 725
 Db 60934 GCTCAAAATTCGTTTGTATTAACATTCGCTGTGTGCTATATCTCTTTGGCAGGAC 60875
 QY 726 tttaactcagggaatgcacttgagctcacaacaagccagcagcagctgctgccaagt 785
 Db 60874 TTTCACTCAGGAATGCACTTGAGTCTCACAACAGGCCAAGCAGCTGCTGCCAAGTG 60815
 QY 786 aagtggtttcttattagagcttggtgattctctatggaattttgctgtgagctttatgaaa 845
 Db 60814 AAGTGGTTTCTTATTAGAGCTGTGATTTCTATGGAATTTTGCCTGTAGCTTTATGAAA 60755
 QY 846 ccttcgagatttcttatttcttcttctgataacagagagcttaataagaagagcagtt 905
 Db 60754 CTTTCGGATTTCTTATATTTCTTTTGTATAACAGAGTCTTAATGAAGAGAGCCAGTT 60695
 QY 906 ggaagtctta 914
 Db 60694 GGAGTCTTA 60686
 RESULT 5
 LOCUS AY059893
 DEFINITION Arabidopsis thaliana histone deacetylase-like protein (At5g22650);
 1172 bp mRNA linear PLN 04-NOV-2001
 ACCESSION AY059893
 VERSION MDJ22.7
 KEYWORDS FLI_CDNA.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1172)
 AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission

JOURNAL

COMMENT

Submitted (22-OCT-2001) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
 Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
 Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..1172

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="v"

/clone="RAFL08-18-E14"

/note="This clone is in pBluescript"

1..1172

/gene="At5g22650; MDJ22.7"

75..995

/gene="At5g22650; MDJ22.7"

/codon_start=1

/product="histone deacetylase-like protein"

/protein_id="AAL24375.1"

/db_xref="GI:16649047"

/translation="MEFWAVTPKNAKTVTPEDSLVHSQSALDCTVKSSESVLVS
 SVDDDEYPAVPAPAPPTAVTANGAGAAVAVKADTKPKAPAEVKPAEKPEDEDES
 DDEDESEDDSEKMDVEDDDSDDEDESEDEEETPKPKPINKKPNPNSVSKT
 PVSGKKAAPAAAPASTPKTEKKKGCHTATPHPAKKGKSPVNAANGSPKSGSGSG
 NNNKKPNSGKQFGGNNKSGNKGKGGRA"

362 a 234 c 285 g 291 t

BASE COUNT

ORIGIN

Query Match 16.6%; Score 156.2; DB 8; Length 1172;
 Best Local Similarity 64.1%; Pred. No. 2.5e-24;
 Matches 252; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

QY 25 tttctcaaccttgattcttagccatggagttctggtgggaattgaattcaatcaggaaag 84

Db 51 TTCTCTCTCTCTCTCTCAACAACAATGCGAGTTCTGGGAGTTGCGGTGACACCAAAAC 110

QY 85 ccagttacagtgactcctgaagaaggcaattcttccacgtttctcaagcaalcgttga 144

Db 111 GCTACTAAGGTGACTCTTCAAGAGACAGACAGCTTGTCCACATTTCTTCAGCCTCACTT 167

QY 145 gaatgtataaacaagaaggagagttgtgctttacatgtataaaggttgggaaccagaac 204

Db 168 GACTGCACAGTCAAAATCTGGAGAAATCTGTGTTTGTAGTGTGACTGTGTGGGGCTAAA 227

QY 205 ttggttctgggaactctatcagctgaggaacatccctcagctttctgtgattgtgattc 264

Db 228 CTTGTTATTGGAACACTTTTCAACAGCAAGTTCCCTCAGATTAGCTTTGATTGTTT 287

QY 265 gacaaggagtttgagcttctccacacttggggaaaaagaaagtgcttacttcttgatgac 324

Db 288 GATAAAGAGCTTTGAGCTTTTCACACAGCGGTACCAAGCAAAATGTTCACTTTTCATTG 347

QY 325 aaaaactcccaacttgagcccaaggcttctctgaggaagaagaagaagaagaaga 384

Db 348 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCCGATGATGAGGATGTTCTCTGAA 407

repeat_region	similar to chromosome V (7873-8427), indirect repeat (copy B, 8428-8440), direct repeat (copy B, 8441-8443)* /note-"Region: Putative transposon" /evidence=not_experimental 7860..7872 /note-"Inverted repeat flanking putative transposon from position 7873 to 8427, copy A" /rpt_type=inverted 7860..8443 /note-"These 586 nucleotides are 97.9% identical to a region of chromosome V clone MBK5 (AB005234), position 81694 to 82277 and encompass the flanking inverted repeats" /note-"Region: Similar to clone MBK5" complement(8428..8440) /note-"inverted repeat flanking putative transposon from position 7873 to 8427, copy B" /rpt_type=inverted 8441..8443 /note-"direct repeat flanking putative transposon from position 7873 to 8427, copy B" /rpt_type=direct 8575..8700 /note-"inverted repeat 1, copy A; copy B located from 9649 to 9775" /rpt_type=inverted 8872..9139 /note-"direct repeat 1, copy A; copy B located from 9789 to 10041" /rpt_type=direct complement(9013..9156) /note-"predicted by Grail" /evidence=not_experimental complement(9446..9741) /note-"predicted by Grail" /evidence=not_experimental complement(9477..9741) /note-"Predicted by MZEF; shows moderately weak similarity to EST F14401" /evidence=not_experimental complement(9649..9775) /note-"Inverted repeat 1, copy B; copy A located from 8575 to 8700" /rpt_type=inverted 9789..10041 /note-"direct repeat 1, copy B; copy A located from 8872 to 9139" /rpt_type=direct complement(9819..10293) /note-"predicted by MZEF; shows moderately weak similarity to EST F14401" /evidence=not_experimental complement(9819..10054) /note-"predicted by Grail" /evidence=not_experimental 12951..13621 /note-"region shares 90.3% identity to chromosome V BAC T21B04 (AF007271) from position 32026 to 32690 on the reverse strand" /note-"Region: Similarity to BAC T21B04" complement(14523..17066) /gene="T32N15.3" /note-"gene encoding hypothetical protein similar to T32N15.4" /evidence=not_experimental complement(join(14523..14849,15006..15264,15584..16038,16325..16643,16706..16771,16969..17066)) /gene="T32N15.3"
misc_feature	/note-"T32N15.3 is most similar to hypothetical protein T32N15.4 (most nucleotide sequence differences are second or third position changes separated by 3 to 5 bases of identity, a hallmark of a multigene family rather than a transposon). T32N15.3 shows moderate similarity to the transposon."
gene	
CDS	

Db	262	TCTGTGATTGGTATCGACAGGAGTTTGAGCTTTCACACTTGGGAAAGGAAGTG	321	
QY	308	tttactttgttgatacaaaaactcccaacattgagcccaagctctattctgaggaagaag	367	
Db	322	TTTACTTTTGGATACAAAACCTCCACATTGAGCCCAAGGCTATTCTGAGGAAGAAG	381	
QY	368	aggaagaagaagaagaagctctgctggaagctgcaagagctgactaaaccaaagg	427	
Db	382	AGGAAGAAGAGGAAGAAGTTCCTGTGGGAATGCTGCCAAGGCTAGCTAAACCAAGG	441	
QY	428	ctaagcctcagaagtgaaagcagctgttgatgatgaagagatgactgtcttctgacg	487	
Db	442	CTAAGCCTCCAGAGTGAAGCAGCTGTGTGATGATGAAGAGATGAGTCTGTCTGAGG	501	
QY	488	gaatgatgaagatgattctgtggtgaggtctctgaggaagaagaagcctacacctaaga	547	
Db	502	GAATGATGAAGATGATTCTGATGCTGGGATCTCTGAGGATCTGAGGAAGAGGCTACACCTAAGA	561	
QY	548	agcctgcatacaagaagaagactaatgaaactaccctaaagcactgtgtcagcaa	607	
Db	562	ACCTGCTATCAAGCAAGAGAGAGCTAATGAAACTACCCCTAAAGCAGCTGTGTACGAA	621	
QY	608	agaagcgaagtagcagttactctcagaagaacagatgagaagaagagggggaagag	667	
Db	622	AGAAGCGAAGTAGCAGTTACTCTCAGAAAACAGATGAGAAGAAAGGGGGAAGG	681	
QY	668	ctgcaaacacagagccaaagtcggcagtcgaagtcattgtgttcacgaagaagactt	727	
Db	682	CTGCAAACAGAGCCCAAGTCGGCAGTCAAGTCTCATGTGTGTCATGCAAGAAGACTT	741	
QY	728	tcaactcaggaatgacactgagctctcacaacaagcgaagcagcctgctgccaagtga	787	
Db	742	TCAACTCAGGAATGACATGAGTCTCACAACAGCCCAAGCAGCCTGTCTCAGAGTGA	801	
QY	788	gtggtttctattagagctgtgattctctatggaatttgcctgtagcttttatgaaacc	847	
Db	802	GTGGTTTCTATTAGAGCTGTGATTCTCTATGGAATTTTGCCTGTAGTCTTTATGAAACC	861	
QY	848	ttggattttctattattttctttttgatacaagaagcttaataaagaagagccggttgg	907	
Db	862	TTGCGATTTCTTATATTTCTTTTGTATAACAAGAGTCTTAATGAAGAGAGCCAGTTGG	921	
QY	908	agctctta 914		
Db	922	AGTCTTA 928		
RESULT	4			
AC002534/c				
LOCUS		101371 bp	DNA	linear
DEFINITION		Arabidopsis thaliana BAC T32N15 from chromosome III near 54 cM,		PLN 21-APR-1999
ACCESSION		complete sequence.		
VERSION		AC002534		
KEYWORDS		AC002534.1 GI:2392762		
SOURCE		HTG.		
ORGANISM		thale cress.		
REFERENCE		Arabidopsis thaliana		
AUTHORS		Fukuyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 101371)		
TITLE		de la Bastide, M.R., Parnell, L.D., Kaplan, N., Gnoj, L., Hameed, A., Schutz, K., Hasegawa, A., Gottesman, T., Shohdy, N., Granat, S., Jensen, K., Johnson, A.F., Lodhi, M., Dedhia, N., Martienssen, R. and McCombie, W.R.		
JOURNAL		A. thaliana BAC T32N15 from chromosome V		
AUTHORS		Unpublished		
TITLE		2 (bases 1 to 101371)		
JOURNAL		McCombie, W.R. and de la Bastide, M.		
AUTHORS		Direct Submission		
TITLE		Submitted (11-SEP-1997) Lita Annenberg Hazen Genome Sequencing		
JOURNAL		Center, Cold Spring Harbor Laboratory, 1 Bungtown Rd., Cold Spring		

REFERENCE		Harbor, NY 11724, USA		
AUTHORS		3 (bases 1 to 101371)		
TITLE		Parnell, L.D.		
JOURNAL		Direct Submission		
REMARK		Submitted (21-APR-1999) Lita Annenberg Hazen Genome Sequencing		
COMMENT		Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724		
		Arabidopsis thaliana BAC T32N15 from chromosome III near 54 cM		
		Sequencing, analysis, and annotation were performed in		
		collaboration with the CSHL/WU/ABI Arabidopsis Sequencing		
		Consortium. Information on physical mapping, YAC and BAC library		
		construction, and a more detailed annotation of the sequence can be		
		viewed at http://www.cshl.org/Arabidopsis/ . Gene prediction programs		
		GRail and MZEF were used in the assembly of coding regions. In		
		certain regions of the BAC these exons are not incorporated into a		
		good model gene and are thus annotated as single exons. BAC T32N15		
		has been mapped to chromosome III near 54 cM and marker m249. A		
		graphical display of the annotation of T32N15 can be found at		
		http://www.cshl.org/Arabidopsis/T32N15-titlepage.html .		
FEATURES		Location/Qualifiers		
source		1..101371		
		/organism="Arabidopsis thaliana"		
		/cultivar="Columbia"		
		/db_xref="taxon:3702"		
		/chromosome="V"		
		/map="near 54 cM"		
		<404..1903		
gene		/gene="T32N15.1"		
		/note="gene encoding hypothetical protein; 5' region of		
		T32N15.1 is not likely to be encoded by BAC T32N15"		
		/evidence="not experimental"		
		join(<404..1327,1413..1559,1616..1903)		
CDS		/gene="T32N15.1"		
		/note="amino-terminus is not likely to be encoded by BAC		
		T32N15; sequence is similar to many entries of A. thaliana		
		genomic sequence, for example MX110, GenBank accession		
		number AB005248"		
		/codon_start=1		
		/evidence="not experimental"		
		/product="hypothetical protein"		
		/protein_id="AAB70026.1"		
		/db_xref="GI:2392763"		
		/translation="VPLGSSVSTASLNCPLDELIVHLSLTAKAASVLSKRW		
		TLFVRNLDPSDIISHPVEGQNMDDVQESPRDFVDRKLAFGQSPINKFSLIVGD		
		KHDDVRDMINTALEMGVSELHICLTSVTRRLRRPNSVFRSTLVKLTGLTNLFIV		
		YFPSDTCLPVKILVLDLSIFDKIKFSNVLIAGCPALDLITDKSPGIPVNVYSKT		
		VKSLSIVKYSDADFPRTVALDTPLNLTLYIYARHRYHGNLPSLVNATLDLHFL		
		ENCDEAPENVTDLMIAVRNVQMLHTSSATEVISCCCKGGLPMFKNLLVFLNGTE		
		RMKVFLLLEHSPNLTKLCSLFLILGLOGVGTDEDEFEIHIPRSNKVMRLRI		
		QCQTENELKHISHFLKMECLQLVQVNFSETIVDSKKVLTEDLMKLPSSASRLTMQ		
		VI"		
		3559..3750		
		/note="function-putative_MITES; region is 96.9% identical		
		to the nucleotide sequence of A. thaliana BAC T7N9,		
		GenBank accession number AC000348; although GRAIL predicts		
		an exon from position 3620 to 3759 and translation of this		
		exon is nearly identical to protein T7N9.14;		
		gene-prediction programs fail to assemble this entire exon		
		into an mRNA sequence"		
		/rpt_type="dispersed"		
		6329..6981		
		/note="function-minisatellite; 23 copies of a 28-bp tandem		
		repeat"		
		/rpt_type="dispersed"		
		/rpt_unit="TTTTCGCCGCAAAAACNTAAACCCGTGA		
		7857..7859		
		/note="function-direct repeat; cognate A of direct repeat		
		flanking putative transposon from position 7873 to 8427"		
		/rpt_type="direct"		
		7857..8443		
		/note="putative non-autonomous transposon; genomic		
		organization is: direct repeat (copy A, position		
		7857-7859), indirect repeat (copy A, 7860-7872), region		

Db 361 GAAGAAGAGGAAGGAAGAAGTCTCTCTGGGAATGTCGAAGGCTGTAGCTAAA 420
 Qy 421 ccaagagctaaagctgcagaagtagcagctgtgtgatgtagagagtagagctgat 480
 Db 421 CCAAGGCTAGGCTGCAAGAGTAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGT 480
 Qy 481 tctacgaaatggaatggaatggaatggaatggaatggaatggaatggaatggaat 540
 Db 481 TCTACGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 540
 Qy 541 cctaaagaagctgcatcaagcaagaagagagagagagagagagagagagagagagag 600
 Db 541 CCAAGAAGGCTGCAATCAAGCAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Qy 601 tcaagcaagaaggggcaagtagcagctgagctgagctgagctgagctgagctgagctgag 660
 Db 601 TCAGCAAGAGGCGAAGTAGAGGCTGCTCTCTCAGAAACAGATGAGCAAGAAAGGG 660
 Qy 661 gaaagagctgcaaacag 720
 Db 661 GCAAGGCTGCAACACAG 720
 Qy 721 aaagcttcaactcagggaag 780
 Db 721 AAGACTTCAACTCAGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 Qy 781 aagtagagctgtct 840
 Db 781 AAGTGAAGTGTCT 840
 Qy 841 tgaaccttcagagatttct 900
 Db 841 TGAACCTTCGAGATTCT 900
 Qy 901 cagttgagagctcta 914
 Db 901 CAGTTGAGAGCTCTTA 914

RESULT 3
 AV065396
 Locus Arabidopsis thaliana putative histone deacetylase (At3g44750) mRNA, PLN 11-DEC-2001
 DEFINITION complete cds.
 ACCESSION AV065396
 VERSION AV065396.1 GI:17529219
 KEYWORDS FLI CDNA.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 977)
 Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,
 Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.,
 Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Nguyen, M.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Full length cDNA of gene At3g44750 (GI:15230496)
 Unpublished
 2 (bases 1 to 977)
 Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,
 Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.,
 Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Nguyen, M.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE JOURNAL
 COMMENT
 FEATURES
 source
 1. .977
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAF1.06-70-K17 (R11479)"
 /note="This clone is in a modified phluescript vector
 (lambda ZAP) as a XhoI/SstI insert.
 ecotype: Columbia"
 1. .977
 /gene="At3g44750"
 1. .62
 /gene="At3g44750"
 63. .800
 /gene="At3g44750"
 /codon_start=1
 /evidence=experimental
 /product="putative histone deacetylase"
 /protein_id="AAL38837.1"
 /db_xref="GI:17529220"
 /translation="MEFVGVKSKPKVTVTPBEGILIHVSASLGECKNKKGEFVPL
 HKVGNQNLVLGTSTENIPQLFCDLVDFKFEELSHITWKGVSFVYGVYKPIEPGY
 SEEEEEEVDPAGNAAKAPKAPKAPKAPKAPKAPKAPKAPKAPKAPKAPKAPKAPKAPK
 EEPKPKPASPKNRANETTPKAPVSAKAKAVPTPKTDEKKKGGKAAQNSPKASAOV
 SCSCKKTNSGNALESHNKKAKHAAK"
 801. .977
 /gene="At3g44750"
 311 a 183 c 231 g 252 t
 BASE COUNT
 ORIGIN
 Query Match 96.4%; Score 905.4; DR 8; Length 977;
 Best Local Similarity 99.9%; Pred. No. 1.3e-189;
 Matches 906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 8 ccgtataaatact 67
 Db 22 CCCTAAAAATCCT 81
 Qy 68 aagttaaatcagaagaag 127
 Db 82 AAGTAAATCAGGAAG 141
 Qy 128 ctgagcagctgag 187
 Db 142 CTCAGGAGATCGCTGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 201
 Qy 188 aggttggaag 247
 Db 202 AGGTGGGAACAG 261
 Qy 248 tctgtgag 307

Direct Submission
 Submitted (04-DEC-2001) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAF1 cDNAs (RAF1 cDNA: 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PCEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAF1 cDNAs: Yamada, K., Banh, J.,
 Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T.,
 Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C.,
 Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PCEC) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PCEC)
 contributed equally to this work as PIs.

LOCUS AF195545 914 bp mRNA linear PLN 01-NOV-2000
 DEFINITION Arabidopsis thaliana putative histone deacetylase (HD2A) mRNA,
 complete cds.
 ACCESSION AF195545
 VERSION AF195545.1 GI:11066134
 KEYWORDS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 REFERENCE 1 (bases 1 to 914)
 AUTHORS Wu, K., Tian, L., Malik, K., Brown, D. and Miki, B.
 TITLE Functional analysis of HD2 histone deacetylase homologs in Arabidopsis thaliana
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 914)
 AUTHORS Wu, K., Tian, L., Malik, K., Brown, D. and Miki, B.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
 FEATURES
 source 1..914
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /clone="211N2"
 /gene="HD2A"
 /gene="HD2A"
 /gene="HD2A"
 /codon_start=1
 /product="putative histone deacetylase"
 /protein_id="AAG28472.1"
 /db_xref="GI:11066135"
 /translation="MBFWGTEVSKGKPVTVPEGLIHVSQASLGCKNKGKGFVPL
 HVKVGONLVLTSTENIPOLCDLVDFDKFEFLSTWKGKSVFYCYKTPNLEPGY
 SESEEEFEVFNAGNAKAVKPKAEVKPAVDDEEDSDGDEDDSDGDESEE
 EETPKPASSKKRNETTPKPVSAKARAVTPORTDEKKGKRAANGSPKASQV
 SCUSCKRTFNSNALESHNKAHAAK"
 BASE COUNT 281 a 171 c 230 g 232 t
 ORIGIN
 Query Match 97.3%; Score 914; DB 8; Length 914;
 Best Local Similarity 100.0%; Pred. No. 1.7e-191;
 Matches 914; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 cagcgctccgtaaaatcctctctttctcaaccttgattcttagccatggattctgg 60
 DB 1 CACGCGTCGGTAAATCTCTCTTTTCTCAACCTTGATTCTTAGCCATGGAGTCTGG 60
 QY 61 gaaattgaagttaaatcaggaagccagttacagtgactcctcgaagaaggcattctatc 120
 DB 61 GGAATTGAAGTTAAATCAGGAAGCCAGTACAGTCACTCTCGAAGAGGCACTTCTTATC 120
 QY 121 caggtttcagcgcctcgttggaagatgtaaaacagaaggaggagattgtgccttta 180
 DB 121 CAGGTTCTCAGGCATCGCTTGGAGATGTAAAAACAAGAGGAGAGTGTGTGCTTTA 180
 QY 181 catgtaaaggttggaaccagaacttggttctgggaactctatcgactgagaacatccct 240
 DB 181 CATGTAAAGGTTGGAGAACAGAACTTGGTTCTGGGAACCTCTATCGACTGAGAACTCCCT 240
 QY 241 cagctttctgtgatttggtatcgcagaaggagtttgagcttctcacacttgggaaaa 300
 DB 241 CAGCTTTTCTGTGATTGCTATTGACAAAGAGCTTTGAGCTTTCTCACACTTGGGAAAA 300
 QY 301 gaaagtgttactttgttgatacaaaactcccaacttggccacacaggctattctgag 360
 DB 301 GGAAGTGTTTACTTTGTGGATACAAAACCTCCCAACATTGAGCCACAAAGGCTATTCTGAG 360
 QY 361 gaagaaggaagaagaagaagttctctggtgaatgctggaagctgccaagctgagctaaa 420
 DB 361 GAAGAAGAGAGAGAGAGAGAGAGTTCTCTGCGGAATGCTGCCAGGCTGTAGCTATA 420
 QY 421 ccaaggttaagcctgcgaaagtagaagccagctgttgatgatgaagagagtagtctgat 480
 DB 421 CCAAGGCTTAAGCCTGCAGAAAGTAGAGCAGCTGTGTGATGATGAAGAGGATGAGTCTGAT 480
 QY 481 tctgaaggaatgagatgaatgatctctgagtgaggattctggaagaagaagaagcctaca 540
 DB 481 TCTGAGGAATGGATGAAGATGATTCTGTATGTGAGGATTTCTGAGGAAGAGAGAGCTTACA 540
 QY 541 cctaaagcctgcacaaaggaagagagctaatgaaactaccctaaagcactgtg 600
 DB 541 CCTAAGAGCCTGCATCAAGCAAGAGAGAGAGCTAATGAACCTACCCCTTAAGCACCCTGTG 600
 QY 601 tcagcaagaagcgaaagtagcagttactctcagaaacacagatgagaagaagaaggg 660
 DB 601 TCAGCAAGAAGCGGAAAGTAGCAGTTACTCTCTCAGAAAACAGATGAGAAGAAGAAGGG 660
 QY 661 gaaaggttgcaaacaccagagcccaagtcggccagtcagtcagtcagtcagtcagtcag 720
 DB 661 GGAAGGCTGCAACACAGAGCCCAAGTCGGCCAGTCAAGTCTCATGTGTTGTTATGCAAG 720
 QY 721 aagacttcaactcaggggaatgcaacttgagtcacacaagggccaagcagcagtcagtc 780
 DB 721 AAGACTTTCACTCAGGGAATGCATTTGAGTCTCACAAGGCGCAAGCAGCCTGCTGCC 780
 QY 781 aagtgaaagtgttcttattagagcttggtattctctatggaatttgcctgtagctcttta 840
 DB 781 AAGTGAAGTGGTTCTTATTAGAGCTTGTGATTCTTATGGAATTTTGCCTGTAGTCTTTA 840
 QY 841 tgaaccttcgaatttcttatttcttcttcttcttcttcttcttcttcttcttcttctt 900
 DB 841 TGAACCTTCGGAATTTCTTATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 900
 QY 901 cagttgagctcttaaaaaaaggggggggggggggggggggggggggggggggggggggg 939
 DB 901 CAGTTGAGCTCTTAAAAAAGGG 939
 RESULT 2
 AF195545

ConCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:56:26 ; Search time 5336.92 Seconds
(without alignments)
3681.902 Million cell updates/sec

Title: US-09-645-337-5

Perfect score: 939

Sequence: 1 cagcgctcgtataatact.....aaaaaaaaagggcgccgc 939

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_st.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_st.*

27: em_sy.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description

1	939	100.0	939	6	AX139018	AX139018 Sequence
2	914	97.3	914	8	AF195545	AF195545 Arabidops
3	905.4	96.4	977	8	AY065396	AY065396 Arabidops
4	249	26.5	101371	8	AC002534	AC002534 Arabidops
5	156.2	16.6	1172	8	AY059893	AY059893 Arabidops
6	156.2	16.6	1187	8	AF195546	AF195546 Arabidops
7	156.2	16.6	1218	6	AX139020	AX139020 Sequence
8	149.8	16.0	918	6	AF044914	AF044914 Arabidops
9	124.2	13.2	1236	6	AF372889	AF372889 Arabidops
10	123.8	13.2	1307	6	AF168371	AF168371 Sequence
11	123.4	13.1	885	8	AF255712	AF255712 Arabidops
12	117.6	12.5	894	8	AF255711	AF255711 Oryza sat
13	109.2	11.6	1124	8	AF384033	AF384033 Zea mays
14	104.2	11.1	1245	6	AF168370	AF168370 Sequence
15	103.8	11.1	912	8	AF254072	AF254072 Zea mays
16	100	10.6	1191	6	AF168369	AF168369 Sequence
17	98.4	10.5	903	8	AF254073	AF254073 Zea mays
18	98.4	10.5	1283	6	AF168368	AF168368 Sequence
19	96.8	10.3	1169	8	ZM082815	ZM082815 Zea mays nu
20	95	10.1	77363	8	AB006699	AB006699 Arabidops
21	85.6	9.1	132323	8	AP003416	AP003416 Oryza sat
22	85.6	9.1	174218	2	AC091071	AC091071 Oryza sat
23	77	8.2	7218	6	I66494	I66494 Sequence 14
24	73.2	7.8	612	8	AF255713	AF255713 Arabidops
25	72	7.7	172224	2	AC087149	AC087149 Mus muscu
26	69.8	7.4	75475	8	AB005235	AB005235 Arabidops
27	69.8	7.4	102897	8	AT17C15	AT17C15 Arabidops
28	69.2	7.4	440	8	PAB296729	PAB296729 Picea abi
29	68.6	7.3	388	8	PAB296740	PAB296740 Picea abi
30	68.2	7.3	20717	2	AC013104	AC013104 Drosophill
31	68.2	7.3	182623	3	AC011071	AC011071 Drosophill
32	68.2	7.3	202008	2	AI611970	AI611970 Mus muscu
33	68.2	7.3	205429	2	AC005506	AC005506 Plasmodiu
34	68.2	7.3	301457	3	AE003512	AE003512 Drosophill
35	67.4	7.2	3720	14	S76368	S76368 ORF 5' of E
36	67.4	7.2	43658	14	HSV3PRGEN	HSV3PRGEN
37	67.4	7.2	112950	14	HSV3PRGEN	HSV3PRGEN
38	66.6	7.1	153477	2	AC006278	AC006278 Plasmodiu
39	66.2	7.1	67970	3	PFMAL1P3	PFMAL1P3 Plasmodiu
40	66	7.0	397	11	G37798	G37798 GARP plasm
41	66	7.0	2384	3	PFAGAR	PFAGAR Plasmodium
42	66	7.0	158548	3	PFMAL3P2	PFMAL3P2 Plasmodiu
43	64.8	6.9	2306	3	DDV18	DDV18 D. discoideu
44	64.6	6.9	2486	8	AF026917	AF026917 Zea mays
45	64.6	6.9	179595	2	AC098110	AC098110 Rat. l. us no

ALIGNMENTS

RESULT	1
AX139018	AX139018
LOCUS	Sequence 5 from Patent EP1094112.
DEFINITION	AX139018
ACCESSION	AX139018.1
VERSION	GI:12474700
KEYWORDS	thale cress.
SOURCE	Arabidopsis thaliana
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 939)
AUTHORS	Wu, K., Miki, B. L., Tian, L. and Brown, D. C.
TITLE	Repressing gene expression in plants
JOURNAL	Patent: EP 1094112-A 5 25-APR-2001;
FEATURES	The Minister of Agriculture and Agri-Food (CA)
source	Location/Qualifiers
	1..939
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
BASE COUNT	296 a 175 c 236 g 232 t
ORIGIN	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 17:56:40 ; Search time 665.94 Seconds
(without alignments)
2420.913 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cagcgccgtaaaatcct.....aaaaaaaaagcgccgc 939

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1980.DAT:*
2: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1981.DAT:*
3: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1982.DAT:*
4: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1983.DAT:*
5: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1984.DAT:*
6: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1985.DAT:*
7: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1986.DAT:*
8: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1987.DAT:*
9: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1988.DAT:*
10: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1989.DAT:*
11: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1990.DAT:*
12: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1991.DAT:*
13: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1992.DAT:*
14: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1993.DAT:*
15: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1994.DAT:*
16: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1995.DAT:*
17: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1996.DAT:*
18: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1997.DAT:*
19: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1998.DAT:*
20: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA1999.DAT:*
21: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA2000.DAT:*
22: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA2001A.DAT:*
23: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA2001B.DAT:*
24: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneeqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	939	100.0	939	22	AAF80352	Nucleotide sequence
2	905.4	96.4	991	21	AA45610	Arabidopsis thaliana
3	886	94.4	1002	21	AA35217	Arabidopsis thaliana
4	860	91.6	940	21	AA36931	Arabidopsis thaliana
5	734.4	78.2	774	21	AA43043	Arabidopsis thaliana
6	156.2	16.6	1147	21	AA41072	Arabidopsis thaliana
7	156.2	16.6	1212	22	AAF80353	Nucleotide sequence
8	124.2	13.2	741	21	AA40930	Arabidopsis thaliana
9	123.8	13.2	1307	20	AA490845	Maize histone deac

10	104.2	11.1	1245	20	AA490844	Maize histone deac
11	100	10.6	1191	20	AA490843	Maize histone deac
12	98.4	10.5	1283	20	AA490842	Maize histone deac
13	82.6	8.8	856	21	AA49658	Arabidopsis thaliana
14	77.8	8.3	879	21	AA49657	Arabidopsis thaliana
15	70	7.5	468	21	AA49656	Arabidopsis thaliana
16	68.2	7.3	303	23	AB124558	Drosophila melanog
17	68.2	7.3	2403	23	AB124558	Drosophila melanog
18	61.6	6.6	4000	18	AA491902	Manose-1-phosphat
19	61	6.5	895	21	AA491902	Manose-1-phosphat
20	60.4	6.4	1260	24	AB199204	Arabidopsis thaliana
21	60	6.4	534	21	AA492224	Mouse ischaemic co
22	59	6.3	475	22	AB442952	Arabidopsis thaliana
23	59	6.3	475	22	AB442952	Human breast cell
24	59	6.3	475	22	AB453373	Human foetal liver
25	59	6.3	475	22	AB453373	Probe #1619 for ge
26	59	6.3	475	22	AA491642	Human brain expres
27	59	6.3	475	22	AA491642	Human bone marrow
28	59	6.3	475	22	AA491680	Human bone marrow
29	59	6.3	475	22	AA491680	Probe #1613 for ge
30	59	6.3	475	22	AA491680	Probe #1666 used t
31	58.8	6.3	3224	18	AA491872	Rat YTS21 gene. R
32	57.8	6.2	2518	20	AA491872	Arabidopsis thaliana
33	57.8	6.2	2518	20	AA491872	Human V3 loop HIV
34	56.8	6.0	3399	17	AA492553	Human nucleolin DM
35	56.6	6.0	3399	17	AA492553	Chicken leucocytot
36	56.6	6.0	3901	18	AA493445	Human p160 CDNA 16
37	56.6	6.0	3901	18	AA493445	Human p160 CDNA 16
38	56.6	6.0	4999	20	AA493445	Murine LOBO genom
39	56.4	6.0	4999	20	AA493445	Murine LOBO genom
40	56.4	6.0	291	23	AA493445	Murine LOBO genom
41	55	5.9	1368	21	AA493445	Plasmodium falci
42	54.8	5.8	372	23	AA493445	DNA encoding novel
43	54.8	5.8	801	23	AA493445	DNA encoding novel
44	54.8	5.8	1072	23	AA493445	DNA encoding novel
45	54.6	5.8	1024	21	AA493445	DNA encoding novel

ALIGNMENTS

RESULT 1	
AAF80352	standard; DNA; 939 BP.
XX	
XX	
AC	AAF80352;
XX	
DT	29-JUN-2001 (first entry)
XX	
DE	Nucleotide sequence of a histone deacetylase designated AthDA2A.
XX	
KW	Histone deacetylase; AthDA2A; RP03; gene expression; transgenic plant;
KW	HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX	
OS	Arabidopsis thaliana.
XX	
FH	Key
FT	Location/Qualifiers
FT	CD5 49..786
FT	/tag= a
FT	/product= "histone deacetylase AthDA2A"
XX	
XX	
PN	CA2316036-A1.
PD	27-FEB-2001.
XX	
PF	24-AUG-2000; 2000CA-2316036.
XX	
PR	27-AUG-1999; 99US-0383971.
XX	
PA	(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX	
PI	Miki B, Brown D, Tian L, Wu K;
XX	

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

QY 128 ctcaagcatcgcttgagaaatgtaaaacaagaaggagatttgcccttaacatgttaa 187
 DB 142 ctcaagcatcgcttgagaaatgtaaaacaagaaggagatttgcccttaacatgttaa 201
 QY 188 aggttgagaaacccaagcttggtcttggaactctacagcttagaacatccctcagctt 247
 DB 202 aggttgagaaacccaagcttggtcttggaactctacagcttagaacatccctcagctt 261
 QY 248 tctgtatttggtatctcgacaagagatttgagcttctcacccttggggaaaggagtg 307
 DB 262 tctgtatttggtatctcgacaagagatttgagcttctcacccttggggaaaggagtg 321
 QY 308 ttactcttggtgatacaaaactcccaacattgagcccaaaagctattct-----g 358
 DB 322 ttactcttggtgatacaaaactcccaacattgagcccaaaagctattcttgaggaagag 381
 QY 359 aggaagaagaagaaagaaagaaagaaagctcctcgctgagatctgcacagctgtagcta 418
 DB 382 aggaagaagaagaaagaaagaaagaaagctcctcgctgagatctgcacagctgtagcta 441
 QY 419 aaccaaaggctaaagctcgagaaagtgaagccagctgttgatgatagaagaagtagctg 478
 DB 442 aaccaaaggctaaagctcgagaaagtgaagccagctgttgatgatagaagaagtagctg 501
 QY 479 attctgacggaatgagtagaagatgattctgattgtagagagattctgagagaagaagccta 538
 DB 502 attctgacggaatgagtagaagatgattctgattgtagagagattctgagagaagaagccta 561
 QY 539 caactaagaagctcgatcaagcaagaagaagctaatgaactacccctaaagccactg 598
 DB 562 caactaagaagctcgatcaagcaagaagaagctaatgaactacccctaaagccactg 621
 QY 599 tctcagaagaagaagcggaagtagcagcttactcctcagaagaacagatgagaagaagaag 658
 DB 622 tctcagaagaagaagcggaagtagcagcttactcctcagaagaacagatgagaagaagaag 681
 QY 659 ggggaaaggctcgaacccaagcccaagaagctcggccagctcaaaatctcagtggttatgta 718
 DB 682 ggggaaaggctcgaacccaagcccaagaagctcggccagctcaaaatctcagtggttatgta 741
 QY 719 agagaagcttcaactcagggaaatgacatgagctctacacacaagaagcccaagccagctgctg 778
 DB 742 agagaagcttcaactcagggaaatgacatgagctctacacacaagaagcccaagccagctgctg 801
 QY 779 ccaagtgaaagtgttctctatagagctgtgattctctatggaatttgcctgtagctt 838
 DB 802 ccaagtgaaagtgttctctatagagctgtgattctctatggaatttgcctgtagctt 861
 QY 839 latgaacacctcgatttctatlatattctcttctgtaacaagaagcttlaatgaagaaga 898
 DB 862 tatgaacacctcgatttctatlatattctcttcttgaataacaagaagcttlaatgaagaaga 921
 QY 899 gccagttgagactcta 914
 DB 922 gccagttgagactcta 937
 RNSUIT 4
 AAC36931
 ID AAC36931 standard; DNA; 940 BP.
 AC AAC36931;
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15573.
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW Protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX

PN EP1033405-A2.
 XX 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135621.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.


```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 91.6%; Score 860; DB 21; Length 940;
 Best Local Similarity 97.7%; Pred. No. 4,9e-213;
 Matches 899; Conservative 0; Mismatches 5; Indels 16; Gaps 2;

```

QY 8 ccgttaaatccctctctcttctcaacctgattcttagccatggagcttgggaattg 67
DB 22 cccataaaatccctctcttctccaaaccttgattcttagccatggagcttgggaattg 81
QY 68 aagttaatcagaaagccagttacagtgactcctgaagaagcattctatccagttt 127
DB 82 aagttaatcagaaagccagttacagtgactcctgaagaagcattctatccagttt 141
QY 128 ctaagcatcgtcttggaagtgttaaaacaagaaggaggttggccttaccatgtaa 187
DB 142 ctaagcatcgtcttggaagtgttaaaacaagaaggaggttggccttaccatgtaa 201
QY 188 aggtttggaaccgaacttggtcttggaactctatcgaactggaacatccccaagtt 247
DB 202 aggtttggaaccgaacttggtcttggaactctatcgaactggaacatccccaagtt 261
QY 248 tctgtgatttgatctgcacaaggagtttgagcttctcacaacttgggaaaggaagt 307
DB 262 tctgtgatttgatctgcacaaggagttt-aggcttctcacaacttgggaaaggaagt 320
QY 308 ttactctgttgatatacaaaatcccaacactgagccacaaggctatcttgagaagaag 367
DB 321 ttactctgttgatatacaaaatcccaacactgagccacaaggctatcttgagaagaag 380
QY 368 aggaagaagaagaagaagcttcctgc-----tgggaatgctgcacaagctg 412
DB 381 aggaagaagaagaagaagcttcctgcgtgcgttactgctgagtggaatgctgcacaagctg 440

```

QY 413 tagtaaaccaaaagcttaagccttcgcaagaatgaaagccagctgtttgatgaaagatg 472
 Db 441 taqtaaaccaaaagcttaaaagcttcgcaagaatgaaagccagctgtttgatgaaagatg 500
 QY 473 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 532
 Db 501 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 560
 QY 533 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 592
 Db 561 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 620
 QY 593 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 652
 Db 621 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 680
 QY 653 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 712
 Db 681 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 740
 QY 713 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 772
 Db 741 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 800
 QY 773 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 832
 Db 801 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 860
 QY 833 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 892
 Db 861 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 920
 QY 893 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 912
 Db 921 agctatattcgaagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 940

PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144633.
 PR 20-JUL-1999; 99US-0144884.

RESULT 5
 AAC43043
 ID AAC43043 standard; DNA: 774 BP.
 AC AAC43043;
 DT 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 37816.
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37816.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX Arabidopsis thaliana.
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.

```
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.2%; Score 734.4; DB 21; Length 774;
Best Local Similarity 99.9%; Pred. No. 1.6e-180;
Matches 735; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 49 atggagttctgggaattgaagttaaatcaggaagccaggttacagtgactcctgaagaa 108
Db 1 atggagttctgggaattgaagttaaatcaggaagccaggttacagtgactcctgaagaa 60
QY 109 ggcattcttaccacgtttctcagggcattcgttggaagataaaacaaagaggagag 168
Db 61 ggcattcttaccacgtttctcagggcattcgttggaagataaaacaaagaggagag 120
QY 169 ttgtgctttacattgataaagggttgggaaccagaaacttgggttggaactctatcgact 228
Db 121 ttgtgctttacattgataaagggttgggaaccagaaacttgggttggaactctatcgact 180
QY 229 gagaacatccctcagctttctgtgatttggttattcgacaagaggtttgagctttctcac 288
Db 181 gagaacatccctcagctttctgtgatttggttattcgacaagaggtttgagctttctcac 240
QY 289 acttggggaagaaagaggtttactttgttgatacaaaactcccaacttgaagccacaa 348
Db 241 acttggggaagaaagaggtttactttgttgatacaaaactcccaacttgaagccacaa 300
QY 349 ggcattctgaggaagaaagaggaaggaaggaaggaaggaaggaaggaaggaaggaag 408
Db 301 ggcattctgaggaagaaagaggaaggaaggaaggaaggaaggaaggaaggaaggaag 360
QY 409 gctgtagctaaacaaagcctgaagcctgagaaagtaagcctgaagcctgtgatgaagag 468
Db 361 gctgtagctaaacaaagcctgaagcctgagaaagtaagcctgtgatgaagag 420
QY 469 gatgagctgattctgacggaattggtgaagatgattcgtgattgaggaattctgagga 528
Db 421 gatgagctgattctgacggaattggtgaagatgattcgtgattgaggaattctgagga 480
QY 529 gaagagcctacacctaaggaagcctgcataagcaagaaagagagcctaataaactacccct 588
Db 481 gaagagcctacacctaaggaagcctgcataagcaagaaagagagcctaataaactacccct 540
QY 589 aaagcactgtgtcagcaaaagagggcgaagtagcaggttagcaggttagcagaaacagatgag 648
Db 541 aaagcactgtgtcagcaaaagagggcgaagtagcaggttagcaggttagcagaaacagatgag 600
QY 649 aagaagaaagggggaagcctgcaaacacagagcccaagtcggcagtcagtcattgt 708
Db 601 aagaagaaagggggaagcctgcaaacacagagcccaagtcggcagtcagtcattgt 660
QY 709 ggttcattgcaagagagctttcaactcaggaattgctgacttgagcttcacaaagggcgaag 768
Db 661 ggttcattgcaagagagctttcaactcaggaattgctgacttgagcttcacaaagggcgaag 720
```

[illegible]

PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149436;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159294;
PR	13-OCT-1999;	9905-0159295;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160980;
PR	22-OCT-1999;	9905-0160981;
PR	22-OCT-1999;	9905-0160989;
PR	25-OCT-1999;	9905-0161404;
PR	25-OCT-1999;	9905-0161405;
PR	25-OCT-1999;	9905-0161406;
PR	26-OCT-1999;	9905-0161359;
PR	26-OCT-1999;	9905-0161360;
PR	26-OCT-1999;	9905-0161361;
PR	28-OCT-1999;	9905-0161920;
PR	28-OCT-1999;	9905-0161921;
PR	28-OCT-1999;	9905-0161992;
PR	28-OCT-1999;	9905-0161993;
PR	29-OCT-1999;	9905-0162142;

[illegible]

RESULTS 7

REF ID: A580353

AAF60333
ID AAF80353 standard: DNA: 1212 BP.

XX AAF80353:

XX
DT 29-JUN-2001 (first entry)

XX DE Nucleotide sequence of a histone deacetylase designated AtHDA2B.

KW Histone deacetylase; AHD2B; RPD3; gene expression; transgenic plant;
KW HDAL; ethylene-responsive phenotype; hypocotyl elongation; ds.

OS Arabidopsis thaliana.

XX	Key	Location/Qualifiers
FH	CDS	61..978
FT		

CA2316036-A1.

PD 27-FEB-2001.

24-AUG-2000: 2000CA-2316036.

XX 27-AUG-1999; 99US-0383971.

PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX PI Miki B, Brown D, Tian L, Wu K:

AA
DR
WPI; 2001-258457/27.

DR P-PSDB; AAB67814.

Methods for regulating gene expression in transgenic plants, e.g. repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl elongation), comprises introducing genes encoding histone deacetylase.

XX
PS
Claim 9: Fig 2B: 91pp: English-

xx The present sequence encodes an Arabidopsis thaliana histone deacetylase
cc designated AtHD2B. The protein is homologous to yeast RPD3 and HDAC.
cc The polynucleotide sequence is used in the method of the invention.
cc The specification describes a method for regulating gene expression in
cc transgenic plants. The method comprises modifying histones by introducing
cc chimeric nucleotide sequences which have regulatory elements in operative
cc association with a gene of interest or with a nucleotide sequence
cc encoding histone deacetylase. The method is useful for regulating the
cc developmental, physiological or biochemical pathway within a plant.

PR	19-JUL-1999;	99US-0144334;
PR	20-JUL-1999;	99US-0144335;
PR	20-JUL-1999;	99US-0144352;
PR	20-JUL-1999;	99US-0144632;
PR	20-JUL-1999;	99US-0144884;
PR	21-JUL-1999;	99US-0144814;
PR	21-JUL-1999;	99US-0145086;
PR	21-JUL-1999;	99US-0145088;
PR	22-JUL-1999;	99US-0145085;
PR	22-JUL-1999;	99US-0145087;
PR	22-JUL-1999;	99US-0145089;
PR	22-JUL-1999;	99US-0145192;
PR	23-JUL-1999;	99US-0145154;
PR	23-JUL-1999;	99US-0145218;
PR	23-JUL-1999;	99US-0145224;
PR	26-JUL-1999;	99US-0145276;
PR	27-JUL-1999;	99US-0145913;
PR	27-JUL-1999;	99US-0145918;
PR	27-JUL-1999;	99US-0145919;
PR	28-JUL-1999;	99US-0145951;
PR	28-AUG-1999;	99US-0146386;
PR	02-AUG-1999;	99US-0146388;
PR	02-AUG-1999;	99US-0146389;
PR	03-AUG-1999;	99US-0147038;
PR	04-AUG-1999;	99US-0147204;
PR	05-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	06-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147303;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	20-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;
PR	23-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	28-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-AUG-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0158293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;

PR	14-OCT-1999;	99US-01593329;
PR	14-OCT-1999;	99US-01593330;
PR	14-OCT-1999;	99US-01593331;
PR	14-OCT-1999;	99US-01596377;
PR	14-OCT-1999;	99US-01596387;
PR	18-OCT-1999;	99US-01595884;
PR	21-OCT-1999;	99US-01607411;
PR	21-OCT-1999;	99US-01607677;
PR	21-OCT-1999;	99US-01607688;
PR	21-OCT-1999;	99US-01607707;
PR	21-OCT-1999;	99US-01608144;
PR	21-OCT-1999;	99US-01608815;
PR	22-OCT-1999;	99US-01609880;
PR	22-OCT-1999;	99US-01609881;
PR	22-OCT-1999;	99US-01609899;
PR	25-OCT-1999;	99US-01614004;
PR	25-OCT-1999;	99US-01614005;
PR	25-OCT-1999;	99US-01614006;
PR	26-OCT-1999;	99US-01613599;
PR	26-OCT-1999;	99US-01613600;
PR	26-OCT-1999;	99US-01613661;
PR	28-OCT-1999;	99US-01619200;
PR	28-OCT-1999;	99US-01619992;
PR	28-OCT-1999;	99US-01619993;
PR	29-OCT-1999;	99US-01621442;

Query Match 13.2%; Score 124.2; DB 21; Length 741;

Best Local Similarity 51.1%; Pred. No. 2.4e-22;
Matches 291; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

Qy	45	agccatggagtcttggggaattgaagttaaatcagaaaagccaggttacagtgactcta	104
Db	114	aacaatggagtcttgggttgtgaagttgaagaatgcgaaccacctcaactagatcctgg	173
Qy	105	agaagcatcttatccaqctttctcaggcatcgottggagaatgtaaaaaacaagaagg	164
Db	174	gtagacagcctcgtgcacatctctcaggttgcttgggagagtagaacaataaac	233
Qy	165	agatttgcctttacatgaaggttgggaaccagaacttgcttctgggaactcttc	224
Db	234	cgaaccaatccaqctttcacgtgactgttgatcggataagcttctcatlggaacgtatc	293
Qy	225	gactgaacaatccctcagcttttctgtgatttggatttcgacaagaaggtttgagcttic	284
Db	294	tcatgagaagttctcagctgtctacggagattgtcttggaaaggaaactttgcgtgc	353
Qy	285	tcacactgggaaaaaggaagtttacttctgttgatacaaaactccccaatctgac	344
Db	354	tcatacttgagaatggggagctttctctctgttacaagttgatgcactctgacc	413
Qy	345	acaaggctattctgaggaagaaggaaaggaaggaaggttctctgctgggaatgctgc	404
Db	414	cgaacctgaggttattgatgatcaacttgagctgctgtttcaaagctgctccgaa	473
Qy	405	caaggcttgacttaaaccaaggctaaacctgcagaagtgaagccagctgttgatgata	464
Db	474	atctgctgcaaacaggtgaacttccagtttcagttgccaaatgaagatgtcaagcccag	533
Qy	465	agagatgactctgattctacggaatggatgaagatgattctgagtggtgaggtattcga	524
Db	534	tgatgcgcgtgacvgtagtgaagaagattctcagatgatgatgagtagaagaaactctgg	593
Qy	525	ggaagaagcctacactaagaagcctgcatacgaagaagaagagagcttaataaactac	584
Db	594	agacgaggaagaagaaggttactgctgaatctgcagctgagaagagagactcatcaga	653
Qy	585	ccctaaagccactgtgtcagcaagaaggg	613
Db	654	cgacqaagaagatdactctctcaugaagg	682

RESULT 9
AAX90845

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0134370.
PR 27-JUL-1999; 99US-0134768.
PR 27-JUL-1999; 99US-0134941.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

RESULT 15
 AAC37832
 ID AAC37832 standard; DNA; 468 bp.
 XX
 AC AAC37832;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18825.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130831.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0137222.
 PR 01-JUN-1999; 99US-0137528.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.

```
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.5%; Score 70; DB 21; Length 468;
Best Local Similarity 57.4%; Pred. No. 2.3e-08;
Matches 124; Conservative 1; Mismatches 91; Indels 0; Gaps 0;

QY 437 cagaagtgaagcagctgttgatgaagagatgaagtctgtattctgcgcggaatgatg 496
    | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 ctgaagagatgatgactctgagaaaggaatggtgttgatgaagatgactcagatgatg 311
    | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 497 aagatgattctggtgagattctgagaaagagcctacacctaaagaagcctgcacat 556
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 acgagggaggagattctgagagagagagagagagagactcctataaagcctgagccaa 371
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 557 caagcaagaagagatgaactaacccctaaagcacctgtgtcagcaagaagcgga 616
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 372 tcaacaagaagaggccaaatgaattctgtatccaracacccctctctggaagaaggcaa 431
QY 617 aagtagcagttactctctcagaaaaacagatgagaaga 652
    || ||||| | || | |||||
Db 432 aaccagcagcagcaccagcttctactctcagagaaga 467
```

Search completed: May 2, 2002, 20:26:19
Job time: 8979 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 16:03:33 ; Search time 100.71 Seconds
(Without alignments)
2290.238 Million cell updates/sec

Title: US-09-645-337-5
Perfect score: 939
Sequence: 1 cagcgctcgtataaatctct.....aaaaaaaaaagggcgccgcg 939

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123.8	13.2	1307	4	US-09-282-305-17
2	104.2	11.1	1245	4	US-09-282-305-15
3	100	10.6	1191	4	US-09-282-305-13
4	98.4	10.5	1283	4	US-09-282-305-11
5	77	8.2	7218	1	US-08-232-463-14
6	57.8	6.2	2518	4	US-09-433-699-3
7	56.6	6.0	3211	2	US-08-574-959A-8
8	56.6	6.0	3211	4	US-09-357-014-8
9	56.6	6.0	3901	2	US-08-574-959A-6
10	56.6	6.0	3901	4	US-09-357-014-6
11	55.6	5.9	1164	4	US-09-538-871-3
12	55.6	5.9	1232	4	US-09-538-871-1
13	52.2	5.6	3489	2	US-08-728-323A-1
14	52.2	5.6	32207	2	US-08-770-379-20
15	52.2	5.6	32207	4	US-08-757-669A-20
16	52.2	5.6	32207	4	US-09-230-371A-20
17	49.4	5.3	2295	1	US-08-375-300-3
18	49.4	5.3	2295	3	US-09-177-431-3
19	49.4	5.3	2295	5	PCT-US95-16930-3
20	49.4	5.3	4080	1	US-08-375-300-1
21	49.4	5.3	4080	3	US-09-177-431-1
22	49.4	5.3	4080	5	PCT-US95-16930-1
23	48.6	5.2	966	2	US-08-766-738-2
24	48.4	5.2	759	1	US-08-466-603-4
25	48.4	5.2	759	1	US-08-314-503A-4
26	48.4	5.2	759	1	US-08-468-066-4
27	48.4	5.2	759	2	US-08-466-717-4

```

28 48.4 5.2 759 3 US-08-466-743-4 Sequence 4, Appli
29 48.4 5.2 759 5 PCT-US95-12414-4 Sequence 4, Appli
30 48.4 5.2 980 1 US-08-466-603-3 Sequence 3, Appli
31 48.4 5.2 980 1 US-08-314-503A-3 Sequence 3, Appli
32 48.4 5.2 980 1 US-08-468-066-3 Sequence 3, Appli
33 48.4 5.2 980 2 US-08-466-717-3 Sequence 3, Appli
34 48.4 5.2 980 3 US-08-466-743-3 Sequence 3, Appli
35 48.4 5.2 980 5 PCT-US95-12414-3 Sequence 3, Appli
36 47 5.0 289 4 US-09-007-005-17 Sequence 17, Appl
37 47 5.0 289 4 US-09-244-796-17 Sequence 17, Appl
38 47 5.0 2277 1 US-08-676-967-2 Sequence 2, Appli
39 47 5.0 2277 1 US-08-676-974-2 Sequence 2, Appli
40 47 5.0 2277 2 US-09-098-487-2 Sequence 2, Appli
41 46.4 4.9 2787 4 US-09-277-565-27 Sequence 27, Appli
42 46 4.9 1276 4 US-09-177-325-2 Sequence 2, Appli
43 46 4.9 1276 4 US-09-411-812A-2 Sequence 2, Appli
44 46 4.9 1276 4 US-09-590-113-2 Sequence 2, Appli
45 45.6 4.9 43795 3 US-08-742-185-101 Sequence 101, App

```

ALIGNMENTS

```

RESULT 1
US-09-282-305-17
: Sequence 17, Application US/09282305
: Patent No. 6287843
: GENERAL INFORMATION:
: APPLICANT: Baldwin, Donald A.
: APPLICANT: Briggs, Steven P.
: APPLICANT: Crane, Virginia C.
: TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
: FILE REFERENCE: 5718-44,
: CURRENT APPLICATION NUMBER: US/09/282,305
: PRIOR FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: 60/080,563
: PRIOR FILING DATE: 1998-04-03
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 17
: LENGTH: 1307
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (87)..(944)
US-09-282-305-17

```

Query Match 13.2%; Score 123.8; DB 4; Length 1307;
Best Local Similarity 54.5%; Pred. No. 1e-24;
Matches 298; Conservative 0; Mismatches 237; Indels 12; Gaps 2;

```

QY 41 tcttagccttgaggttcttgggaattgaagttaataacagaaagccaggttaccgtgactc 100
Db 79 tcttcccaatggaggttcttggggggaagagtgagccagggagccaggttcttgcgaag 138
QY 101 ctgaagaagcattcttaccacgttcttcagcagcgtcttgagagatgtataaaacaaga 160
Db 139 ttggtgaggttgggttatccaccctttccacaggtcgccttagggaa---ccaaagaaag 195
QY 161 agggagaggttggcctttacatgtataaagttgggaaccagaacttggttctgggaactc 220
Db 196 cgaagtgaagatgccattgtgtctgctcaaaattgtataagaactagtgcttgggaactc 255
QY 221 tatcgactgagaacatccctcagctttctgtgatttgatttcgacaaggaatttggagc 280
Db 256 tatcaattgagaagcattccaaatctctgtgatttgatttcgataaagatttggat 315
QY 281 ttctcacacttgggaaggaagtgatttactcttcttgatatacaaaactcccaacatcg 340
Db 316 tatcacacaattcaagacagatagtttctctgtggtctacaagtcacctgttctc 375

```


Qy 507 tgaatgagatctctgaggaagaagcctacacctaagaagcctgacataagcaaga 566
Db 1226 RRR 1167
Qy 567 gagagctaataaaccctccctlaagcacctgtgtcagcaaaagagcgaaatagcagt 626
Db 1166 RRR 1107
Qy 627 tactcctcagaaacagatgaggaagaagggggggaagggcgtgcaaac 675
Db 1106 RRR 1058

RESULT 6
US-09-433-699-3
; Sequence 3, Application US/09433699B
; Patent No. 6165786
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
; FILE REFERENCE: RTS-0109
; CURRENT APPLICATION NUMBER: US/09433.699B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (112)..(2235)
US-09-433-699-3

Query Match 6.28; Score 57.8; DB 4; Length 2518;
Best Local Similarity 49.58; Pred. No. 1.9e-06;
Matches 149; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

Qy 347 aagctattctgaggaagaagaagaaggaaggaagttcctgctgggaatgctgcca 406
Db 551 aagaggagatgagcagtgaggagatgagaggatgagcaggaagagatgagatg 610
Qy 407 aggtgtagctaaacaaaggcctaagcctgcaggaagtgaaagcagctgttgatgaag 466
Db 611 aagatgaattgaaccagcagatgaagcagcagctgctgcccctgcclcaagagatg 670
Qy 467 aggatgagctgattctgacgaagatgagatgaagatgattctgaggtgaggtctgagg 526
Db 671 aggcagtgaggatgacgaagatgagatgagatgagatgagatgagatgagatgactctg 730
Qy 527 aagaagagcctacacctaaggaagcctgcacacgaagaagaagaagcctaaactacc 586
Db 731 aagaagaagctatgagactacacacagcagcagcagcagcagcagcagcagcagcag 790
Qy 587 ctgaagcagctgtgtcagcaagaagaagcagcagcagcagcagcagcagcagcagcag 646
Db 791 tgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 850
Qy 647 a 647
Db 851 a 851

RESULT 7
US-08-574-959A-8
; Sequence 8, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jounq, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08574.959A
FILING DATE: 19-DEC-95
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: DFN-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3211 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 439..3157
US-08-574-959A-8

Query Match 6.08; Score 56.6; DB 2; Length 3211;
Best Local Similarity 57.7%; Pred. No. 4.4e-06;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 358 gaggaagaagaagaagaagaagaagttcctgctggaagcagcagcagcagcagcagcagc 417
Db 2407 GAGGAAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2466
Qy 418 aaaccacaaagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 477
Db 2467 GAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2526
Qy 478 gattctgacgggaatgagatgagatgattctgaggtgaggtctgaggtgaggtgaggt 532
Db 2527 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2581

RESULT 8
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jounq, Ratna K. Vadlamudi
; APPLICANT: and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible


```
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3847
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-357-014-6

Query Match          6.0%; Score 56.6; DB 4; Length 3901;
Best Local Similarity 57.7%; Pred. No. 4.9e-06;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 358 gaggaagaaggaggaaggaaggaaggtctctctggtggaatgctgcgaaggctgtagct 417
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3097 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3156

QY 418 aaacccaaaggctaaagcctgcagaaagtgcagccagctgttgatgatgaagaggaagct 477
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3157 CAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3216

QY 478 gattctcagaatggaatgaagatgattctgattgattgattgattgattgattgattg 532
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3217 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3271

RESULT 11
US-09-538-871-3
; Sequence 3, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark
; APPLICANT: Szebenl, Atilla
; TITLE OF INVENTION: METHODS FOR STABILIZATION AND
; TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
; FILE REFERENCE: 13121.000202
; CURRENT APPLICATION NUMBER: US/09/538,871
; EARLIER FILING DATE: 2000-03-30
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
; OTHER INFORMATION: synthetic construct
US-09-538-871-3

Query Match          5.9%; Score 55.6; DB 4; Length 1164;
Best Local Similarity 50.8%; Pred. No. 5.2e-06;
Matches 133; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 292 tggggaaaaaggagtggtttactttgttgatatacaaaactcccaacttgagccacaaggc 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 tgtggtctggccttgacataaagtgacacacacacacacacacacacacacacacac 444

QY 352 tattctgagaagaaggaggaaggaaggttctctgctgggaatgctgcgaaggct 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 gagtcagaagatgaagatgaggaagatgataaaactcttaggcattgctggaagagatct 504

QY 412 gtagctaaacaaaggctgaagcctgcagaaagtgaagccagctgttgatgatgaagagat 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 gctccccggaggtgtaacaaagtcacagaaaaaagtaaaacttgatgaagatgatgat 564

QY 472 gagtcattctgacggaalgaatgaagatgattctgattgaggaattctgaggaagaa 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 gaggatgatgaagatgatgaagatgatgaagatgatgaagatgatgaagatgatgaag 624
```

```
QY 532 gagcctacacacctaagaagcctg 553
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 gaaactgaagaaaaaggttccag 646

RESULT 12
US-09-538-871-1
; Sequence 1, Application US/09538871
; Patent No. 6358718
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark
; APPLICANT: Szebenl, Atilla
; TITLE OF INVENTION: METHODS FOR STABILIZATION AND
; TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS H23
; FILE REFERENCE: 13121.000202
; CURRENT APPLICATION NUMBER: US/09/538,871
; EARLIER FILING DATE: 2000-03-30
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1232
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
; OTHER INFORMATION: synthetic construct
US-09-538-871-1

Query Match          5.9%; Score 55.6; DB 4; Length 1232;
Best Local Similarity 50.8%; Pred. No. 5.3e-06;
Matches 133; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 292 tggggaaaaaggagtggtttactttgttgatatacaaaactcccaacttgagccacaaggc 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 tgtggtctggccttgacataaagtgacacacacacacacacacacacacacacacac 415

QY 352 tattctgagaagaaggaggaaggaaggttctctgctgggaatgctgcgaaggct 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 gagtcagaagatgaagatgaaggaagatgataaaactcttaggcattgctggaagagatct 475

QY 412 gtagctaaacaaaggctgaagcctgcagaaagtgaagccagctgttgatgatgaagagat 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 gctccccggaggtgtaacaaagtcacagaaaaaagtaaaacttgatgaagatgatgat 535

QY 472 gagtcattctgacggaalgaatgaagatgattctgattgaggaattctgaggaagaa 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 gaggatgatgaagatgatgaagatgatgaagatgatgaagatgatgaagatgatgaag 595

QY 532 gagcctacacacctaagaagcctg 553
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 gaaactgaagaaaaaggttccag 617

RESULT 13
US-08-728-323A-1
; Sequence 1, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 15:42:36 ; Search time 3265.32 Seconds
(without alignments)
3881.286 Million cell updates/sec

Title: US-09-645-337-5

Perfect score: 939

Sequence: 1 cacqgcctcgtataaatcct.....aaaaaaaaagggcgccgc 939

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	636	67.7	636	9 AV544477	AV544477 AV544477
2	451.2	48.1	534	9 AI995198	AI995198 701502609
3	367.6	39.1	434	10 N37851	N37851 19078 Lambd
4	156.2	16.6	422	10 BE523968	BE523968 M4A9STM
5	156.2	16.6	728	10 BE844690	BE844690 AD01A077
6	149.8	16.0	504	10 T45874	T45874 9137 Lambd
7	146.8	15.6	543	10 BF279194	BF279194 GA_EB003
8	145.8	15.5	378	9 AV551556	AV551556 AV551556
9	141.6	15.1	759	10 BG131117	BG131117 EST464009
10	141.6	15.1	875	10 BF275467	BF275467 GA_EB002
11	141	15.0	327	9 AV549937	AV549937 AV549937
12	140.2	14.9	511	10 B1130765	B1130765 G110P49r
13	137.8	14.7	865	10 BG445639	BG445639 GA_Ea002
14	136.2	14.5	517	10 BF278973	BF278973 GA_Eb003
15	136	14.5	715	10 BF270691	BF270691 GA_Eb000
16	135.6	14.4	382	10 C99809	C99809 C99809 YAC
17	131.2	14.0	560	10 B1786647	B1786647 sal51f09.

18	130.4	13.9	537	10 BM177387	BM177387 esj80q01.
19	129.8	13.8	539	9 AW906751	AW906751 EST342874
20	129.2	13.8	595	9 AW929159	AW929159 EST337947
21	128.6	13.7	479	10 B1924066	B1924066 EST543955
22	128.6	13.7	504	10 B134445	B134445 EST467337
23	128.6	13.7	537	9 AW615938	AW615938 EST325304
24	128.6	13.7	542	9 AW929950	AW929950 EST354220
25	128.6	13.7	552	10 B134979	B134979 EST467871
26	128.6	13.7	555	10 BE449829	BE449829 EST361267
27	128.6	13.7	589	10 B1934377	B1934377 EST554266
28	128.6	13.7	590	10 B1930352	B1930352 EST550241
29	128.6	13.7	611	10 B1924005	B1924005 EST543894
30	128.6	13.7	668	10 B1927751	B1927751 EST547640
31	128	13.6	674	10 BG595734	BG595734 EST494412
32	127.6	13.6	435	9 AW429116	AW429116 EST306572
33	127.6	13.6	579	10 BE343018	BE343018 EST395862
34	127.6	13.6	591	10 BE343072	BE343072 EST395916
35	127.6	13.6	661	10 BF460278	BF460278 073G03 Ma
36	127	13.5	453	10 BF187099	BF187099 EST443386
37	127	13.5	467	9 A1776222	A1776222 EST257322
38	127	13.5	467	10 BE342419	BE342419 EST395263
39	127	13.5	473	10 BE922565	BE922565 EST426334
40	127	13.5	548	10 BG594032	BG594032 EST492710
41	127	13.5	578	10 B1433232	B1433232 EST535993
42	126.6	13.5	387	9 AW039158	AW039158 EST281393
43	125.8	13.4	686	10 BG594317	BG594317 EST492995
44	125.6	13.4	581	10 BG889339	BG889339 EST515190
45	125.6	13.4	635	10 BG597002	BG597002 EST495680

ALIGNMENTS

RESULT 1

AV544477/c

LOCUS

AV544477

DEFINITION

CDNA clone R245C06F 3', mRNA sequence.

ACCESSION

AV544477

VERSION

EST

KEYWORDS

EST

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

20363093

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yama 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1. 636

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="R245C06F"

/columbia.lib="Arabidopsis thaliana roots Columbia"

/tissue-type="roots"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT

ORIGIN

147 a 167 c 121 g 201 t

Query Match

67.7% Score 636; DB 9; Length 636;

Best Local Similarity 100.0%; Pred. No. 3.7e-112;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 ctatcgactgagacatccctcagctttctgtgattgttgatttcacaaaggagtttgag 279
 |||||
 Db 636 CTATCGACTGAGAACATCCCTCAGCTTTCTGTGATTGGTATTGTCACAGGAGTTTGAG 577

QY 280 ctltctcacacttggggaaagaaagtgatttacttctgttgatttcacaaactcccaact 339
 |||||
 Db 576 CTTTCTCACACTTGGGAAAGAGAGTGTATTACTTTCTGGATACAAAACCTCCCAACATT 517

QY 340 gacccaaagctatctcgaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaat 399
 |||||
 Db 516 GAGCCACAGGCTATTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457

QY 400 gctgccaagctgtagctaaacaaagctaaagctcctgcagaagctgaagccagctgtgat 459
 |||||
 Db 456 GCTGCCAAGCTGTAGTAAACCAAGGCTAAGCTGCGAGAGTGAAGCCAGCTGTGTAT 397

QY 460 gatgaagagatgagctgaltctgcaggaatggatgaagatgattctgattgaggaat 519
 |||||
 Db 396 GATCAAGAGGATGAGTCTGATTTCTGACGGAATGGATCAAGATGATTCTGATGCTGAGGAT 337

QY 520 tctgaagaaagagctacacctaagaagcctgcataagcaagaagaagaaagaaagaaatgaa 579
 |||||
 Db 336 TCTGAGGAAGAGAGGCTACACTAAGAGGCTGCTATCAAGCAAGAGAGAGCTAATGAA 277

QY 580 actaccctaaagcactgtgtcagcaagaaagcgaagcgaagtagcagttactctcagaaa 639
 |||||
 Db 276 ACTACCCCTAAAGCACCCTGTGTCAGCAAGAGGCGCAAGAGTAGCAGTACTCTCAGAAA 217

QY 640 acagatgaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 699
 |||||
 Db 216 ACAGATGAGAAGAAAGAAAGGGGAAAGGCTGCAACAGAGCCCAAGTCCGCCAGTCAA 157

QY 700 qtcctatgtgttcacgaagaaagccttcaactcaggaatcagcattgagttcacaac 759
 |||||
 Db 156 GTCTCATGTGTTCTATGCAAGAGAGCTTCACTCAGGAATGCATTCAGTCTCACAAC 97

QY 760 aagcccaagcagctgctgccaagtgaaagtggtttcttattagagctgtgattctatg 819
 |||||
 Db 96 AAGGCCAAGCAGCTGCTGCCAAGTGAAGTGTCTTATTAGAGCTTGTGATTCTATG 37

QY 820 gaatttgcctgtagcttctatgaacacctcgatt 855
 |||||
 Db 36 GAATTTTGCTGTAGCTTTATGAAACCTTCGGATT 1

RESULT 2
 AI995198
 LOCUS 701502609 A. thaliana, Ohio State clone set Arabidopsis thaliana
 DEFINITION cDNA clone 701502609, mRNA sequence.
 ACCESSION AI995198
 VERSION AI995198.1 GI:5842103
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 534)
 Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
 Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
 Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
 Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T.,
 Pollock, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L.,
 Turner, C., Krizorian, S., Elder, L., and Hanson, D.
 Arabidopsis thaliana Gene Expression Microarray
 Unpublished (1999)
 Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte
 Pharmaceuticals, Inc.

FEATURES
 source 1..534
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="701502609"
 /clone_lib="A. thaliana, Ohio State clone set"
 /note="cDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."
 BASE COUNT 155 a 95 c 143 g 141 t
 ORIGIN

Query Match 48.1%; Score 451.2; DB 9; Length 534;
 Best Local Similarity 98.1%; Pred. No. 9.8e-77;
 Matches 530; Conservative 0; Mismatches 3; Indels 7; Gaps 7;

QY 11 taaaaatcctctctttctcaaccttgcattcttagccatgagttctgggaattgag 70
 |||||
 Db 1 TAAAAATCCTCTCTCTTTCTCAACCTTGATCTTACCCATGGAGTCTGGGAAATTGAAG 60

QY 71 ttaaatcaggaagccagtlacaglgactcctgaaagagcattcttalccacgtttctc 130
 |||||
 Db 61 TTAATATCAGGAAGCCAGTTACAGTGACTCCCTGAAAGAGGATCTTATCCACGTTTCTC 120

QY 131 aggcctcgttgagaatgttaaaacaagaagggagagttgtgcctttacatgtaagg 190
 |||||
 Db 121 AGGCATCTGCTTGGAGAAATGTAAAAACAAGAGGGAGAGTTTGTGCTTTTACATGTAAAG 180

QY 191 ttgggaacacgaactgtgtctgggaactctctacgtcagcagcaccctcagctttct 250
 |||||
 Db 181 TTGGG-ACCAGAACCTTGGTCTGGGAACCTCTATCCACTGACAACTCCCTCAGCTTTTCT 239

QY 251 gtgatttgatttcacaaagagtttgagctttctcacacttggggaagaaagagtttt 310
 |||||
 Db 240 GTGATTGTGATTTCGCAAGGAGTTTGAGCTTCTCACACTTGGGGAAGAGTGT 299

QY 311 acctgttggtacaaaactcccaacttgccacaaggctattcttgagaagaagg 370
 |||||
 Db 300 ACTTTGTTGGATACAAAACCTCCCAACTTGAGCCACAAAGGCTATTCTGAGGAAGAGG 359

QY 371 aagaagagaaagagttctcctgctgggaatgtcgcacagctgtagctaaacaaaggcta 430
 |||||
 Db 360 AAGAAGAGGAACAAGTCTCTGCTGGGAATGCTGCCAACAGGCTAGCTAAACCAAGGCTA 419

QY 431 agcctgcagaagtgagccagcgtlgtalgaagagagtgagctgattctgacggaa 490
 |||||
 Db 420 AG-CTTCAGAGAGTG-AGCCAGCTGTGTATGATG-AGAGGATCAGTCTG-TTCTGACGCA 475

QY 491 tggatgaagatgattctgattgagagattctgagaagaagagagcctacacctaagaagc 550
 |||||
 Db 476 TGGATGAGC-TGATTCTGATGTTGAGGATTCTGAGGAGAGAG-CTACACCTAAGAGC 533

RESULT 3
 N37851
 LOCUS 19078 Lambda-PRL2 Arabidopsis thaliana cDNA clone 211N2T7, mRNA
 DEFINITION sequence.
 ACCESSION N37851
 VERSION N37851.1 GI:1158993
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 434)
 Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh
 L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Reitel

TITLE
 JOURNAL
 COMMENT

TITLE Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 COMMENT Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@ibm.cl.msu.edu
 Seq primer: T7 dye primer.
 Location/Qualifiers
 1. .434
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="211N27"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA."

FEATURES

source
 117 a 78 c 106 g 116 t 17 others
 BASE COUNT
 ORIGIN
 Query Match 39.1%; Score 367.6; DB 10; Length 434;
 Best Local Similarity 92.6%; Pred. No. 1.1e-60;
 Matches 399; Conservative 0; Mismatches 29; Indels 3; Gaps 2;
 QY 11 taaatactctctcttttctcaaccttgattcttagccatggagttctggggaattgaag 70
 Db 1 TAAANATCCTCTCTTTTNTCAACCTTCAATCTTAGCCATGGAGTCTCGGGGAATTGAAG 60
 QY 71 ttaatacagaagaagccagttacagttacagttacagttacagttacagttacagttctc 130
 Db 61 TTAANATCAGGAAGCCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 120
 QY 131 aggcacgtctggaggaatgtataaacaagaaggaggtttgtgacctttacatgtataaagg 190
 Db 121 AGGCATCGCTGGAGATGTATAAACAAGAAGGAGAGATTTGTNCCCTTACATGTAAAGG 180
 QY 191 ttgggaaccagaacttgggtctgggaactctatcgactgagaacatccctcagctttct 250
 Db 181 TTGGGAACCAAGAACTTGGTCTGGAACTCTATCGACTGAGACATCCCTCAGCTTTNT 240
 QY 251 gtgatttggtattgacagaaggtttgagttcttctcaacttggggaaggaaggtttt 310
 Db 241 GTGATTGGTATTGCAAGAAGGTTTTCAGCTTCTCACATTTGGGGAAAAGGAAGTTTT 300
 QY 311 acttgttgatataaactcccaacttgagccacaaggctattctgagaagaag--a 368
 Db 301 ACTTTGNGGTAGCAAAACTCCCAACATTGAGCCACAAGGCTATTCTCAGGAAGGAGNG 360
 QY 369 ggaagaagaagaag-aagttctgtctgggaagtgtgccaaggtgtgagctaaaccaaagg 427
 Db 361 GGAGNAGGAGGNAGNAAGTTCCTNCTGGGGATTCTGCAAGNCTTTNGGCNAACCCAANG 420
 QY 428 ctaagcctgca 438
 Db 421 GTTAGCCTTCA 431

RESULT 4

BE523968
 LOCUS BE523968 422 bp mRNA linear EST 19-MAR-2001
 M44A9STM Arabidopsis developing seed Arabidopsis thaliana cDNA
 clone M44A9 5', mRNA sequence.
 ACCESSION BE523968
 VERSION BE523968
 KEYWORDS EST;
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1 (bases 1 to 422)
 AUTHORS White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
 Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Benning,C.
 TITLE A new set of Arabidopsis expressed sequence tags from developing
 seeds. The metabolic pathway from carbohydrates to seed oil
 JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)
 MEDLINE 20567808
 COMMENT Contact: Benning, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 , USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TEL: 6142929371.
 Location/Qualifiers
 1. .422
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="M44A9"
 /clone_lib="Arabidopsis developing seed"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoII"
 BASE COUNT 109 a 99 c 92 g 122 t
 ORIGIN

Query Match 16.6%; Score 156.2; DB 10; Length 422;
 Best Local Similarity 64.1%; Pred. NO. 3.3e-20;
 Matches 252; Conservative 0; Mismatches 138; Indels 3; Gaps 1;
 QY 25 ttcttcaaccttgattcttagccatggagttctggggaattgaaatataacaggaaag 84
 Db 31 TTCTCTCTTCTCTGTTCAACAACAATGGAGTTCTGGGAGTTCGGGTGACACCAAAAAC 90
 QY 85 ccagttcacagttacgtcctcgaagaagcattcttccacgtttctcagcagctgttga 144
 Db 91 GCTACTAAGGTGACTCTCTGAAGAAGACAGCCTTGTCCACATTTCTCAGGCTTCACTT 147
 QY 145 gaatgtaaaacaagaaggaggtttgtgacctttacatgtaaaagttgggaaccagaac 204
 Db 148 GACTGCACAGTGAATCTGGAGAAATCTGTGTTTTCAGTGTGTGCTGTGGTGGGGTAA 207
 QY 205 ttgttctgggaactctatcgactgagaacatccctcagcttttctgtgatttgggtatc 264
 Db 208 CTTGTTATTGGAACTTTCACAAGACAAGTTCCTCCCTCAGATTAGCTTTGATTGGTTT 267
 QY 265 gacaagaggtttgagctttctcacacttgggggaaggaaggtttcactttgttgatac 324
 Db 268 GATAAAGAGTTTGAGCTTTTTCACACAGCGGTACCAAGAAATGTTTATTTCATTGCTAC 327
 QY 325 aaaactcccaacattgagccacaaggctattctgagaagaagaaggaaggaagaa 384

```

Db 328 AATCCCAACATCGAGGAGTACTTCACTAGTTCCGATGATGAGGATGTTCTCGAA 387
QY 385 gtctcgtggaatgctgccaggtgtagct 417
Db 388 GCTTCTCTGCTCTGCTGCCCTACTGCTGTACT 420

RESULT 5
LOCUS BE844690 728 bp mRNA linear EST 17-MAY-2001
DEFINITION AD01A07T7 AD A. thaliana (Col-0 gll) library enriched for
salt-induced transcripts; 10-14 day seedlings; 4h 160mM NaCl stress
Arabidopsis thaliana cDNA clone AD01A07 similar to (AF044914)
putative histone deacetylase, mRNA sequence.
ACCESSION BE844690
VERSION BE844690.1 GI:10277068
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 728)
REFERENCE Gong,Z., Kolwa,H., Cushman,M.A., Ray,A., Bufford,D., Kore-Eda,S.,
Matsumoto,T.K., Zhu,J., Cushman,J.C., Bressan,R.A. and Hasegawa
,P.M.
Genes that are uniquely stress regulated in salt overly sensitive
(sos) mutants
Plant Physiol. 126 (1), 363-375 (2001)
21249177
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Plate: AD01 row: A column: 7
Seq primer: T7
POLYA-No. Location/Qualifiers
1..728
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia gll"
/db_xref="taxon:3702"
/clone="AD01A07"
/clone_lib="AD A. thaliana (Col-0 gll) library enriched
for salt-induced transcripts; 10-14 day seedlings; 4h
160mM NaCl stress"
/tissue_type="mixed leaf and root"
/dev_stage="10-14 day old"
/lab_host="Escherichia coli, TOP10F"
/note="Organ: whole seedlings; Vector: pT-Adv (CLONTECH);
Site_1: EcoRI; Site_2: EcoRI; AD Arabidopsis thaliana
library enriched for salt-induced transcripts from
wildtype (Col-0 gll) 10-14 day-old seedlings treated with
160 mM NaCl for 4 hours."
BASE COUNT 219 a 151 c 179 g 179 t
ORIGIN

Query Match 16.6%; Score 156.2; DB 10; Length 728;
Best Local Similarity 64.1%; Pred. No. 2.8e-20;
Matches 252; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

QY 25 ttctcaacttgattcttagccatgagttctgaggaattgaattaaatcaggaag 84
Db 29 TTTCTTCTCTGCTCTGCTCAACAACAAATGGAGTTCTGGGAGTTCGGTGACACCAAAAC 88
QY 85 ccagttacagtactcctggaaggcattcttatcacagtttctcaggcatcgcttga 144
Db 89 GCTACTAAGGTGACTCTGAGAGAGACAGCGCTTGTCCACATTTCTCAGGCTTCATT--- 145

```

```

QY 145 gaatgaaaaacaagaagggagagttgtgacctttacatgtaaaagttlqggaaaccagaac 204
Db 146 GACTGCACAGTGAATCTGGAGAAATCTGTGGTTTTCAGTGTGACGTCTGGTGGGCTAAA 205
QY 205 ttggttctgggaactctatcgactgagaaacatccctcagcttcttctgtgattgttatcc 264
Db 206 CTGTGTTATTGGACACTTTTCAAGACAAAGTTCCCTCAGATTAGCTTTGATTGGTTTTT 265
QY 265 gacaagagtttgagctttctcacacttggggaagaagagtgatttacttttttgatcac 324
Db 266 GATAAGAGTTTGGAGCTTTTCACACACCGGTACCAACAGCAAAATCTTCATTTCATGGCTAC 325
QY 325 aaaaactcccaacttgagccacaagagctattctgaggaagaagagaaagagaaagaa 384
Db 326 AAATCCCCCAACATCGAGCAGGATGACTTCACTAGTTCCGATGATCAGGATCTTCTCGAA 385
QY 385 gtctcgtggaatgctgccaggtgtagct 417
Db 386 GCTGTCTCTGCTCTGCTGCCCTACTGCTGTACT 418

RESULT 6
LOCUS T45874 504 bp mRNA linear EST 09-JAN-1998
DEFINITION 9137 Lambda-PRL2 Arabidopsis thaliana cDNA clone 133F1577, mRNA
sequence.
ACCESSION T45874
VERSION T45874.1 GI:934127
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 504)
REFERENCE Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomas,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On Apr 14, 1993 this sequence version replaced gi:638462.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.ci.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1..504
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="133F1577"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT 131 a 113 c 109 g 142 t 9 others
ORIGIN

```

```

BASE COUNT      149 a   121 c   121 g   152 t
ORIGIN

Query Match      15.6%; Score 146.8; DB 10; Length 543;
Best Local Similarity 60.3%; Pred. No. 2e-18;
Matches 288; Conservative 0; Mismatches 172; Indels 18; Gaps 2;

QY 18 cctctcttttttcaaccttgattcttagcctggaggtcttggggaattgaagttlaaatc 77
   || || || || || || || || || || || || || || || || || || || || ||
Db 58 CCCTTCCTTTTCTTCGTAGTAATCATCCATGGAGTCTGGGTATTGAGTTAAGTC 117

QY 78 aggaagccaggttacagttactcctcgaaggaaggaattcttccacgtttctcaggcatc 137
   || || || || || || || || || || || || || || || || || || || || ||
Db 118 TGGACAGCCTATTAAAGCAGATCCTGGTGTCTAACTAGCTTATCCATCTTCCCAGGCTTC 177

QY 138 gcttgagaagtgtaaaaaagaagggagaggttgtgctttacatataaaggttggaaa 197
   || || || || || || || || || || || || || || || || || || || || ||
Db 178 ACTTGGAGAGTCCAAGAAC---AAAGCAGAAATCCGTTCCACTCTATGTGAATGTTGATGG 234

QY 198 ccagaacttggttctgggaactctatcgactgagaacatccctcagcttttctgtgatt 257
   || || || || || || || || || || || || || || || || || || || || ||
Db 235 GAAGAAACTTATCTGGGAACCTTCCCACGAGGTGGCCCTCAATATCTTTTGATT 294

QY 258 ggtattcgacaagagtttgagctttctcacacttggggaaaaggaagtttactttgt 317
   || || || || || || || || || || || || || || || || || || || || ||
Db 295 AGTATTTGACGAAGAGTTTGAGCTCTCCACAACTCGAAAACGGGAGCGTATATTTCCCT 354

QY 318 tggatacaaaactcccacacttgagccacagaaggtattct-----gagga 362
   || || || || || || || || || || || || || || || || || || || || ||
Db 355 CGGTTACAAAACTTTCATACCAGAGGAAGGTGATGATTTTGGTATGTGTAGCGAAGAAGA 414

QY 363 agaagaggaagaagagaagaagttctctctggaatgctgcgaagctgtagctaaacc 422
   || || || || || || || || || || || || || || || || || || || || ||
Db 415 AAGCAGAGGAGGAGGAAGGAGGAGCTTCTGTTGCTGCTGCTGCTGCTGCGATTAATGG 474

QY 423 aaaggctaaagcctcgagaagtgaaagccagctgttgatgaagaagatgagtctgat 480
   || || || || || || || || || || || || || || || || || || || || ||
Db 475 AAAAGCTAAGCCAGATGGAACAGCAAGGCTAATGCTGGGAACCTGATGCTGTT 532

RESULT 8
AV551556
LOCUS
DEFINITION
  AV551556 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
  cDNA clone RZ128C12R 5', mRNA sequence.
ACCESSION
  AV551556
VERSION
  AV551556.1 GI:8722969
KEYWORDS
  EST.
SOURCE
  thale cress.
ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosoids II; Brassicales; Brassicaceae; Arabidopsids.
  1 (bases 1 to 378)
AUTHORS
  Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE
  A large scale analysis of cDNA in Arabidopsis thaliana: Generation
  of 12,028 non-redundant expressed sequence tags from normalized and
  size-selected cDNA libraries
JOURNAL
  DNA Res. 7, 175-180 (2000)
MEDLINE
  20363093
COMMENT
  Contact: Erika Asamizu
  The First Laboratory for Plant Gene Research
  Kazusa DNA Research Institute
  Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
  Email: asamizuk@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
  Location/Qualifiers
  1..378
  /organism="Arabidopsis thaliana"
  /strain="Columbia"
  /db_xref="taxon:3702"
  /clone_lib="RZ128C12R"
  /clone_lib="Arabidopsis thaliana roots Columbia"
FEATURES
  source

```



```

ORGANISM  Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE  1 (bases 1 to 875)
AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
          D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE     An integrated analysis of the genetics, development, and evolution
          of the cotton fiber
JOURNAL   Unpublished (2000)
COMMENT   Contact: Wing RA
          Clemson University Genomics Institute
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Seq primer: TAATACGACCTCACATATAGGG
          High quality sequence stop: 630.
FEATURES   source
            1..875
             /organism="Gossypium arboreum"
             /strain="AKA"
             /cultivar="8400"
             /db_xref="taxon:29729"
             /clone_lib="GA_EB0024C01f"
             /tissue_type="Fibers isolated from bolls harvested 7-10
             dpa"
             /lab_host="E. coli"
             /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 265 a 191 c 203 g 216 t
ORIGIN
Query Match      15.18; Score 141.6; DB 10; Length 875;
Best Local Similarity 58.18; Pred. No. 1.7e-17;
Matches 293; Conservative 0; Mismatches 199; Indels 12; Gaps 2;

QY  30 tcaacctgattcttagccatggagttctggggaattgaattaaatcaggagaagccagt 89
    || || || || || || || || || || || || || || || || || || || ||
DB  79 TCTTTTCATCATCTCATAAATGGAGTCTCGGGTATTGAAGTTAAAGCTGGACAGCCTAT 138

QY  90 tacagtgactcctaagaagcattcttaccagttcttcaggcagctgttgagaagt 149
    || || || || || || || || || || || || || || || || || || || ||
DB  139 TAAACGGACCGCTGTGGCAGTCATGTCATCTTTCCAGGCAACACTTGGTGAGT- 197

QY  150 taaaaacaagaaggagagatttgcctttacatgtaaaagttgggaaccagaacttgg 209
    || || || || || || || || || || || || || || || || || || || ||
DB  198 --CAAAGAACAAGACAGAGTCCGTTCCTCTCTATGTGAATGTTAATGGGAAGAACTCGT 255

QY  210 tctgggaactctatcagctgagaacatccctcagctttctgtgtatttgcgacaa 269
    || || || || || || || || || || || || || || || || || || || ||
DB  256 TCTCGAACACTTTCCACACAGAAATGCCCGCAATTGCTTTGATTTAGTTATTGACCA 315

QY  270 ggaagttgagctttctcacacttggggaagaagtggtttactttgttgatacaaac 329
    || || || || || || || || || || || || || || || || || || || ||
DB  316 AGGCTTTGAGCTCTCTCACAACTGGAAAAATGGGAGCGTCTATTTCCTTGGTTACAGAAC 375

QY  330 tcccaacattgacca-----caaggctattctgaggagaagaaggaagga 380
    || || || || || || || || || || || || || || || || || || || ||
DB  376 TTTTGTCCAGAGGAAGCGCTGTGATGATGAATTTGATAGTCAGGAAGAAAGCTAGCGAGGA 435

QY  381 agaaagttctctggtggaatgctgcaaggctgtagctaaacaaagcgctaaagcctgcaga 440
    || || || || || || || || || || || || || || || || || || || ||
DB  436 TGAGGAACCTTCTGTGGCTGTGTAACGGAAAGCTTAAGACAGATGCAAAAACGGCTAA 495

QY  441 agtgaagcagctgttgatgatgaagagatgagttctgattctgcaggaatggatgaaga 500
    || || || || || || || || || || || || || || || || || || || ||
DB  496 TGCTGGGAAACCTGATGCTGTAAACAAAGCTGTCAAGATAACTGAACCAAGTGATAACAA 555

QY  501 tgattctgatggtgaggattctga 524
    | ||| || || || || || || || || || || || || || || || || || ||

```

```

Db  556 GAAAAATGAGGACCATGATGATGA 579

RESULT  11
AV549937
LOCUS   AV549937
DEFINITION  Arabidopsis thaliana roots Columbia Arabidopsis thaliana
          cDNA clone RZ106a10R 5', mRNA sequence.
ACCESSION  AV549937
VERSION    AV549937.1
KEYWORDS   EST.
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE  1 (bases 1 to 327)
AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
          of 12,028 non-redundant expressed sequence tags from normalized and
          size-selected cDNA libraries
JOURNAL   DNA Res. 7, 175-180 (2000)
MEDLINE   20363093
COMMENT   Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES   source
            1..327
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZ106a10R"
             /clone_lib="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"
BASE COUNT  88 a 74 c 68 g 97 t
ORIGIN
Query Match      15.08; Score 141; DB 9; Length 327;
Best Local Similarity 67.5%; Pred. No. 2.9e-17;
Matches 214; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

QY  25 ttttctcaaccttgattcttagccatggagttctggggaattgaattaaatcaggaaag 84
    || || || || || || || || || || || || || || || || || || || ||
DB  14 TTTCTCTTCTCTCGTTCAACAACAATGGAGTTCTGGGAGTTGCGGTGACACCAAAAAAC 73

QY  85 ccagttacagtgactcctgaagaaggcattctatccacgtttctcagcgatcgcttga 144
    || || || || || || || || || || || || || || || || || || || ||
DB  74 GCTACTAAGTGACTCTCTGAAGAGACAGCGCTGTGTCACATTTCTCAGGCTTCACTT- 130

QY  145 gaatgtaaaaacaagaaggagagtttgcctttacatgtataaagggttgggaaccagaac 204
    || || || || || || || || || || || || || || || || || || || ||
DB  131 GACTGCACAGTGAATACTGGAGAATCTGTGTTTGGAGTGTGACTGTTGGTGGGCTAAA 190

QY  205 ttggttctgggaactctatcagatgagaacatccctcagcttttctgtgatttggattc 264
    || || || || || || || || || || || || || || || || || || || ||
DB  191 CTTGTTATTGGAACACTTTCAACAACAAGTTCCCTCAGATTAGCTTTGATTGGTTTTT 250

QY  265 gacaaggatttgagctttctcacacttggggaagaagagagtttactttgttgatcac 324
    || || || || || || || || || || || || || || || || || || || ||
DB  251 GATAAAGAGTTTGAGCTTTTACACAGCGGTACCAACAGCAATGTTTCATTTCATTGGCTAC 310

QY  325 aaactcccaacattga 341
    || || || || || || || || || || || || || || || || || || || ||
DB  311 AAATCCCAACATCGA 327

RESULT  12
BI130765

```

LOCUS B1130765 511 bp mRNA linear EST 31-DEC-2001
 DEFINITION G11049Y Populus cambium cDNA library Populus tremula x Populus
 tremuloides cDNA, mRNA sequence.
 ACCESSION B1130765
 VERSION B1130765.1 GI:18014736
 KEYWORDS EST.
 SOURCE Populus tremula x Populus tremuloides.
 ORGANISM Populus tremula x Populus tremuloides.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
 REFERENCE 1 (bases 1 to 511)
 AUTHORS Hertzberg, M., Aspeborg, H., Erlandsson, R., Bjorkbacka, H., Hiltunen
 T., Karlsson, J., Teeri, T., Gustafsson, P., Bahlerao, R., Jansson, S.,
 Nilsson, O., Sundberg, B., Nilsson, P., Uhlen, M., Sandberg, G. and
 Lundberg, J.
 TITLE Gene expression in Populus
 JOURNAL Unpublished (2001)
 COMMENT Contact: Erlandsson R
 Department of Biotechnology
 Royal Institute of Technology
 Teknikringen 30, Stockholm S-10044, Sweden
 Tel: 46 8 790 8287
 Fax: 46 8 245452
 Email: riker@biochem.kth.se.
 FEATURES
 source
 1. 511
 Location/Qualifiers
 /organism="Populus tremula x Populus tremuloides"
 /db_xref="taxon:47664"
 /clone_lib="Populus cambium cDNA library"
 /note="Organ: cambium"
 BASE COUNT 137 a 123 c 100 g 151 t
 ORIGIN

Query Match 14.9%; Score 140.2; DB 10; Length 511;
 Best Local Similarity 64.1%; Pred. No. 3.7e-17;
 Matches 211; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
 QY 47 ccattgagttctgggaattgaatttaatacaggaaagccagttacagtgactccctgaag 106
 Db 155 CCATGGAGTTCGGGTGTTGAAGTCAAGGCTGGTGAGCCTCTTAAGTGAACCTAAAG 214
 QY 107 aaggcatcttaccagttcttcaggcatcgtctgggaagtgtataaaacagaaggag 166
 Db 215 ATTTCTATATGATACATCTTCTTCTCAGGAGCCCTTGCTCAGTCTTCAAGAGGGAACG 274
 QY 167 agtttgctttacatgtataaaggttgggaaccagaaacttggtctcggaaactctatga 226
 Db 275 AATCAGTTCCTCTTCTTAAAGTTGGATGAAAGAACCTTGTTGGGAACCTATCTC 334
 QY 227 ctgagaacatccctcagctttcttgatgttggtattcgacaaggagtttgaagctttctc 286
 Db 335 CTGAAAGATCCCTCAGTCTGCTCTTGATTTAGTTTGTGAGACAGAGTTCGAGCTCTC 394
 QY 287 acacttgggaagaaggagtgattactcttctgttgatacaaaaactcccaacttagccac 346
 Db 395 ACAACTGGAAAAGGGAGGTGTTTCTCTGTGCTACCAAGCTGCTATTCTCTGAAAATG 454
 QY 347 aaggctattctgaggaagaagaagaa 375
 Db 455 ATTCTGATTTTTCGTGATGCGAGATGAA 483

RESULT 13
 RG445639
 LOCUS RG445639 865 bp mRNA linear EST 15-MAR-2001
 DEFINITION GA_Ea0029A01f Gossypium arboreum 7-10 dpa fiber library Gossypium
 arboreum cDNA clone GA_Ea0029A01f, mRNA sequence.
 ACCESSION RG445639
 VERSION BG445639.1 GI:13355291
 KEYWORDS EST.
 SOURCE Gossypium arboreum.

ORGANISM Gossypium arboreum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
 REFERENCE 1 (bases 1 to 865)
 AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry
 D., Wood, T.C., Leslie, A. and Wilkins, T.A.
 TITLE An integrated analysis of the genetics, development, and evolution
 JOURNAL of the cotton fiber
 COMMENT Unpublished (2000)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 High quality sequence start: 5
 High quality sequence stop: 751.
 FEATURES
 source
 1. 865
 Location/Qualifiers
 /organism="Gossypium arboreum"
 /strain="AKA"
 /cultivar="8400"
 /db_xref="taxon:29729"
 /clone_lib="GA_Ea0029A01f"
 /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
 /tissue_type="Fibers isolated from bolls harvested 7-10
 dpa"
 /lab_host="F. coli"
 /note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 251 a 170 c 254 g 189 t 1 others
 ORIGIN

Query Match 14.7%; Score 137.8; DB 10; Length 865;
 Best Local Similarity 53.1%; Pred. No. 9.1e-17;
 Matches 339; Conservative 0; Mismatches 293; Indels 6; Gaps 2;
 QY 30 tcaaccttattcttagccatgaggttcgggaattgaagttaaatcaggaaagccagt 89
 Db 79 TCTTTTCATCTCTCAATAATGGAGTTCGGGTATTTGAAGTTAAAGCTGGACGCTAT 138
 QY 90 tacagtgactctgaagaaggcattcttccacgtttctcaggaatcagcttggagaag 149
 Db 139 TAAACGGACCTGGTCCAGTCATGTCATCTTCCAGGCACACTTGGTGAGT- 197
 QY 150 taaaaacaaggaggaggtttgtgctttacatgtataagggttgggaaccagaaacttgt 209
 Db 198 --CAAGAACAAGCAGAGTCGGTCTCTCTATGTGTAATGTTAATGGGAAGAACTCGT 255
 QY 210 tctggaaactctcgaactgagaacalcctcctgcttcttctgtatllggtatlcgaaa 269
 Db 256 TCTCCGAAACACTTTCACCACTGGGCAATTTGCTGTTTGTATTTAGTATTTGAGCA 315
 QY 270 ggaatttgagcttctccacacttggggaaggaaagagtggttacttlttqgatacaaac 329
 Db 316 AGGCTTTGAGCTCTCCACACTGGGAAATATGGAGCGTCTATTTCTTGGTTAGAAC 375
 QY 330 tcccaacattgagccacaagggtattcttgaggaagaaggaggaagaagaagtctcc 389
 Db 376 TTTTGCCAGAG--GAAGGCTCTGATGATGAAATTTGATGAGGAAGAAATAGCA 432
 QY 390 tgcagggaatctgcgaaggtctgagctaaacaaaggctlaagcctgcgaagtagaagcc 449
 Db 433 GGATGAGGAACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
 QY 450 agctgttgatgatgaaggagatgagctctgattctcagcgaatggatgaagatgattcga 509
 Db 493 TAATGCTGGGAACCTGATGCTGGGAACAACTGTCAAGATATCACTGAACCAAGTGAGTA 552
 QY 510 tggtaggattctgaggaagaagagcctcacactaagaagcctgcatcraagcaagaagag 569

QY 78 aggaagccaggttacagtactctcgaagaaggcattcttattccacggtttctcaggcattc 137
 Db 138 TAGACGCTATTAAAGCAGATCCTGGTGCTAACTAGCTTATCCATCTTTCCAGGCTTC 197
 QY 138 gcttgagaaatgtaaaacaagaaggagaggttctgacctttacatgtaaaaggttgaggaa 197
 Db 198 ACTTGGAGAGTCCAGAAC---AAAGCAGAAATCCGTTCCACTCTATGTGAATGTTGATGG 254
 QY 198 ccadaacttggttcttggaactctatcgactgagaacatccctcagccttttctcgtgaltt 257
 Db 255 GAAGAAACTTATTCTGGGAACCTTTCCACAGAGAGTTGCCCTCAATTATCTTTTGATTT 314
 QY 258 ggtattcgaagaaggtttgagctttctcacttgggggaaaggaggtgttactttgt 317
 Db 315 AGTATTGACGAGAGAGTTTGAGCTCTCCCACTGGAACAGGGAGCGTATATTTCCT 374
 QY 318 tqgatacaaaactcccaacatgagcccaaggcctatctcigaggaagaagaaaga 377
 Db 375 CCGTTACAAAACCTTTCATACCAGAGGAGGTGATGATTTTGGTATGTCTACCGAAGAAGA 434

Search completed: May 2, 2002, 18:43:44
 Job time: 10868 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:14:52 ; Search time 144.83 Seconds
(without alignments)
187.897 Million cell updates/sec

Title: US-09-645-337-6

Perfect score: 1279

Sequence: 1 MEFWGLGVKSGKPVTVTPER.....TFNSGNALESHNKAHAAK 245

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1279	100.0	245	21	Arabidopsis thalia
2	1279	100.0	245	22	Amino acid sequenc
3	1274	99.6	257	21	Arabidopsis thalia
4	1261.5	98.6	248	21	Arabidopsis thalia
5	606	47.4	305	21	Arabidopsis thalia
6	594	46.4	305	22	Amino acid sequenc
7	513.5	40.1	311	20	Maize histone deac
8	504	39.4	98	21	Arabidopsis thalia
9	499	39.0	302	20	Maize histone deac
10	487	38.1	305	20	Maize histone deac
11	446	34.9	285	20	Maize histone deac

12	405	31.7	126	21	Arabidopsis thalia
13	374.5	29.3	208	21	Arabidopsis thalia
14	306	23.9	203	21	Arabidopsis thalia
15	255.5	20.0	155	21	Arabidopsis thalia
16	241	18.8	194	21	Arabidopsis thalia
17	230	18.0	181	21	Arabidopsis thalia
18	224	17.5	412	20	S. frugiperda immu
19	203.5	15.9	103	21	Arabidopsis thalia
20	183.5	14.3	147	21	Arabidopsis thalia
21	175.5	13.7	349	22	Drosophila melanog
22	155	12.1	208	22	Novel human diagno
23	155	12.1	277	22	Novel human diagno
24	155	12.1	317	21	Human Nucleophosmi
25	155	12.1	334	22	Novel human diagno
26	153	12.0	536	22	Human protein sequ
27	153	12.0	707	16	Human nucleolin.
28	153	12.0	707	20	Human V3 loop HIV
29	153	12.0	707	22	Human nucleolin.
30	150	11.7	330	22	Novel human diagno
31	150	11.7	338	22	Novel human diagno
32	147.5	11.5	362	22	Arabidopsis thalia
33	140.5	11.0	268	21	Arabidopsis thalia
34	140.5	11.0	268	21	Arabidopsis thalia
35	140.5	11.0	310	21	Arabidopsis thalia
36	140.5	11.0	310	21	Arabidopsis thalia
37	140.5	11.0	450	21	Arabidopsis thalia
38	140.5	11.0	484	21	Arabidopsis thalia
39	140.5	11.0	590	21	Arabidopsis thalia
40	140.5	11.0	612	21	Arabidopsis thalia
41	140.5	11.0	622	21	Arabidopsis thalia
42	140.5	11.0	649	21	Arabidopsis thalia
43	140.5	11.0	1881	21	Streptococcus pneu
44	140	10.9	244	20	Human secreted pro
45	139.5	10.9	411	20	Yeast immunophilin

ALIGNMENTS

RESULT 1

AAG38221

ID AAG38221 standard; Protein; 245 AA.

AC AAG38221;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 47121.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 23-MAR-1999; 99US-0123548.

XX 25-MAR-1999; 99US-0125788.

XX 29-MAR-1999; 99US-0126264.

XX 01-APR-1999; 99US-0126785.

XX 06-APR-1999; 99US-0127462.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

PR	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	99US-01322048.	PR	21-JUL-1999;	99US-0145086.
PR	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	99US-0132863.	PR	22-JUL-1999;	99US-0145145.
PR	99US-0132867.	PR	23-JUL-1999;	99US-0145218.
PR	99US-0134256.	PR	23-JUL-1999;	99US-0145224.
PR	99US-0134218.	PR	23-JUL-1999;	99US-0145276.
PR	99US-0134219.	PR	26-JUL-1999;	99US-0145913.
PR	99US-0134221.	PR	27-JUL-1999;	99US-0145918.
PR	99US-0134370.	PR	27-JUL-1999;	99US-0145919.
PR	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	99US-0135124.	PR	28-JUL-1999;	99US-0146386.
PR	99US-0135353.	PR	02-AUG-1999;	99US-0146386.
PR	99US-0135629.	PR	02-AUG-1999;	99US-0146388.
PR	99US-0136021.	PR	02-AUG-1999;	99US-0146389.
PR	99US-01366021.	PR	03-AUG-1999;	99US-0147038.
PR	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	99US-0137328.	PR	05-AUG-1999;	99US-0147260.
PR	99US-0137502.	PR	05-AUG-1999;	99US-0147303.
PR	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	99US-0138094.	PR	06-AUG-1999;	99US-0147493.
PR	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	99US-0138847.	PR	09-AUG-1999;	99US-0147935.
PR	99US-0139119.	PR	10-AUG-1999;	99US-0148171.
PR	99US-0139452.	PR	11-AUG-1999;	99US-0148319.
PR	99US-0139453.	PR	12-AUG-1999;	99US-0148341.
PR	99US-0139492.	PR	13-AUG-1999;	99US-0148565.
PR	99US-0139454.	PR	13-AUG-1999;	99US-0148684.
PR	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	99US-0139456.	PR	17-AUG-1999;	99US-0149175.
PR	99US-0139463.	PR	18-AUG-1999;	99US-0149226.
PR	99US-0139750.	PR	20-AUG-1999;	99US-0149722.
PR	99US-0139763.	PR	20-AUG-1999;	99US-0149723.
PR	99US-0139817.	PR	20-AUG-1999;	99US-0149929.
PR	99US-0139899.	PR	23-AUG-1999;	99US-0149902.
PR	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	99US-0139462.	PR	23-AUG-1999;	99US-0150566.
PR	99US-0139463.	PR	25-AUG-1999;	99US-0150884.
PR	99US-0139750.	PR	26-AUG-1999;	99US-0151065.
PR	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	99US-0139899.	PR	27-AUG-1999;	99US-0151303.
PR	99US-0140353.	PR	30-AUG-1999;	99US-0151438.
PR	99US-0140354.	PR	31-AUG-1999;	99US-0151930.
PR	99US-0140695.	PR	01-SEP-1999;	99US-0152363.
PR	99US-0140823.	PR	07-SEP-1999;	99US-0153070.
PR	99US-0140991.	PR	10-SEP-1999;	99US-0153758.
PR	99US-0141287.	PR	13-SEP-1999;	99US-0153758.
PR	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	99US-0142977.	PR	29-SEP-1999;	99US-0156596.

PI Miki B, Brown D, Tian L, Wu K;
XX
XX WPI: 2001-258457/27.
DR N-PSDB; AAF80352.
XX
XX Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT
XX
XX Claim 10; Fig 2A; 9lpp; English.
PS
XX The present sequence represents Arabidopsis thaliana histone deacetylase
XX designated AtHDA2A. The protein is homologous to yeast RPD3 and HDAC1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
XX Sequence 245 AA;
SQ

Query Match 100.0%; Score 1279; DB 22; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.1e-107;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFWGIEVSKGKPVTVPEEGILIHVSOASIGECKNKKGEFVPLRVKVGNONLVIGTLST 60
DB 1 mefwgievskgkpvtlvpeegilihvsgasigecknkkgefvpilhkvgnqlvlgtlst 60
QY 61 ENIPOLFCDLVDFKEFELSHTWGKGSYFVGYKTPNIEPGQYSBEEBEEVVPAGNAAK 120
DB 61 enipqlfcdlvdfkefelshwtgkgsyfvgyktpniepgqysbeebbeevvpagnaa 120
QY 121 AVAKPKAPAEVKPAVDEDESDSGDDEDDSGDESEEEPTPKPASSKKRANETTP 180
DB 121 avakpkapaevkpavdeedesdsgdedsdsgdeeeptpkpasskkranettp 180
QY 181 KAPVSAAKKAVVTPQKTDEKKKGKAANOSPKSASOVSCGSKKTTNSGNALESHNAK 240
DB 181 kapvsakkavvtpqktdekkkgkaanspkasqsvscgskkttinsgnaleshnak 240
QY 241 HAAAK 245
DB 241 haaak 245

RESULT 3
AAG31484
ID AAG31484 standard; Protein; 257 AA.
XX
XX AAG31484;
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 37817.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX

[illegible]

```
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match          99.6%; Score 1274; DB 21; Length 257;
Best Local Similarity 100.0%; Pred. No. 6.3e-107;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFWGLVSKGKPVTVTPBERGILLHVSQASLGECRKNKGFEFVPLHVYKGNQNLVGLTST 60
Db 1 mefwglvskgkpvrtvtpbergillhvsqaslgcknkgfefvplhvkgnqnlvlgltst 60

QY 61 ENIPQLFCDLVDFKELSHTWKSGVYFVGKTPNIEPOGYSSEEEEEEEVPAGNAAK 120
Db 61 enipqlfcdlvdfkelshtwkgsvyfvgyktpniepogysseeeeeevpagnaak 120

QY 121 AVAKPKAKPAEVKPAVDDEEDSDQMDRDSGDSEEEPTPKKPASSKKRARETTP 180
Db 121 avakpkakpaevkpavddeedssdqmdrdsdgedseeeptpkpasskkranettp 180

QY 181 KAPYSAKKAKVATPQKTDKKKGGKAAQSPKASQVSCGCKKTFNSGNALSHKAK 240
Db 181 kapyssakakvavtpqtktdkkkggkaanspkasqvsqgscckktnfsgnalseshnak 240

QY 241 HAAA 244
Db 241 haaa 244
```

RESULT 4

AAG10894

ID AAG10894 standard; Protein; 248 AA.

XX AAG10894;

AC AAG10894;

XX 17-OCT-2000 (first entry)

DT Arabidopsis thaliana protein fragment SEQ ID NO: 9401.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9401.

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9401.

```
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
```


RESULT 5
AAC26185
ID AAG26185 standard; Protein; 305 AA.
XX AC AAG26185;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30547.
XX KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

```

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.4%; Score 606; DB 21; Length 305;
Best Local Similarity 50.3%; Pred. No. 1.5e-46;
Matches 150; Conservative 24; Mismatches 66; Indels 58; Gaps 13;

QY 1 MEFNGIEVKSCKPVTVPTEGILLIHVSQASLGECKNKKGEFVPLHVKGONLVLTST 60
Db 1 mefwgavtpkntkvtpeedslvhisqasl-dctvksgeesvsvltvvggkvlvgtlsq 59

QY 61 ENIPOLFCDLVDFKEFELSHTWGKSVYFVGKTPNIEPOGY-SEEEEEEEVPA---- 115
Db 60 dkfpqisldlvkefslshgqkanvnlgykspnieqddftssddedvpeavpapt 119

QY 116 -----GNAKAV-----AKPKAPAEVKPAVD-----DEDESDSD-----GM--D 149
Db 120 avtangnaqaavvkadskpkapaeekpaekpesdeeddesdeseeddskgmdvd 179

```

```

QY 150 EDDSDG---EDS---EEEEPTPKKPKAS-SKKRANETTPKAPVSAKKAHVAV-----TPOKTD 199
Db 180 eddsdddecdsdeceetpkpkiprncsvsktpvsgkakkpaaapastpqkte 239

QY 200 EKKKGCKA-----ANQSPKASQVSCG-SCKKTFNSGNAL.ESHNKAKH 241
Db 240 ekkkgghtatphakkkgkspvnanqspggsgggnnnkpkpsqkqfgggnkqan 297

RESULT 6
AAB67814
ID AAB67814 standard; Protein; 305 AA.
XX
AC AAB67814;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated AHDA2B.
XX
KW Histone deacetylase; AHDA2B; RPD3; gene expression; transgenic plant;
KW HDAL; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
CA2316036-A1.
XX
27-FEB-2001.
XX
24-AUG-2000; 2000CA-2316036.
XX
27-AUG-1999; 99US-0383971.
XX
(MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
MIKI B, Brown D, Tian L, Wu K;
XX
WPI: 2001-258457/27.
DR N-PSDB; AAF80353.
XX
Methods for regulating gene expression in transgenic plants, e.g.
repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
elongation), comprises introducing genes encoding histone deacetylase
.
Claim 10; Fig 2B; 91pp; English.
XX
The present sequence represents Arabidopsis thaliana histone deacetylase
designated AHDA2B. The protein is homologous to yeast RPD3 and HDAL.
The polynucleotide sequence is used in the method of the invention.
The specification describes a method for regulating gene expression in
transgenic plants. The method comprises modifying histones by introducing
chimeric nucleotide sequences which have regulatory elements in operative
association with a gene of interest or with a nucleotide sequence the
encoding histone deacetylase. The method is useful for regulating the
developmental, physiological or biochemical pathway within a plant,
particularly for repressing ethylene-responsive phenotypes
(e.g. inhibition of hypocotyl elongation). The method is also useful
as a functional test for identifying a phenotype associated with
perturbing a gene. The histone deacetylase genes are useful for
altering the development of an organism.
XX
Sequence 305 AA;
QY
Query Match 46.4%; Score 594; DB 22; Length 305;
Best Local Similarity 50.7%; Pred. No. 1.8e-45;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

QY 1 MEFNGIEVKSCKPVTVPTEGILLIHVSQASLGECKNKKGEFVPLHVKGONLVLTST 60
Db 1 mefwgavtpkntkvtpeedslvhisqasl-dctvksgeesvsvltvvggkvlvgtlsq 59

QY 61 ENIPOLFCDLVDFKEFELSHTWGKSVYFVGKTPNIEPOGY-SEEEEEEEVPA---- 115

```


PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```

PR 29-OCT-1999; 99US-0162142.
Query Match 39.4%; Score 504; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 5.5e-38;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 MDEDDSDGDESEEEPTPKPKPASSKKRANETTPKAPVSAKAKAVAVTPQKTDKKGKGA 207
DDB 1 mdeodsdgdeeeptpkpkpasskkranettpkapsakakavavtpqtkdekkkgga 60
QY 208 ANQSPKASQVSCGCKKTFNSGNALSHNKAHAAK 245
DDB 61 anqspkasqvsqscskttfnsnalsnhnkakhaaak 98

RESULT 9
AAY28803
ID AAY28803 standard; protein; 302 AA.
AC AAY28803;
XX 13-JAN-2000 (first entry)
DE Maize histone deacetylase-7.
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
OS Zea mays.
XX WO9951731-A2.
PN 02-APR-1999; 99WO-US07370.
XX 03-APR-1998; 98US-0080563.
XX (PION-) PIONEER HI-BRED INT INC.
XX Baldwin DA, Briggs SP, Crane VC;
XX WPI; 1999-611038/52.
XX N-PSDB; AAX90843.
XX New deacetylase genes, used for producing transgenic plants which have
XX increased disease resistance -
XX Claim 1; Page 76-77; 87pp; English.
XX The present sequence is maize histone deacetylase encoded by HD cDNA
XX belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
XX modifications, may be localised to promoters targeted for repression by
XX other proteins that associate with HD and specifically bind regulatory
XX elements in promoter DNA. The HD nucleotide sequence can be used for
XX producing transgenic plants with increased disease resistance.
XX Additionally, compositions find use in screening for toxins that affect
XX pathogenicity and in determining which disease response promoters are
XX regulated by histone deacetylase.
SQ Sequence 302 AA;

Query Match 39.0%; Score 499; DB 20; Length 302;
Best Local Similarity 40.5%; Pred. No. 6.5e-37;
Matches 124; Conservative 36; Mismatches 74; Indels 72; Gaps 11;

QY 1 MEFWGIKVGKPVTVTPTEGILIHVSQASLGECKNKKGFVPLRVKVGNONVLGTLS 60
DDB 1 mefwgiwkvkpvstvkcpvgvfhlsqaalge--skksdnalmyvkiddqkkgaitls 58

```

```

QY 61 ENIPQLFCDLVFDKPEFELSHTWGKGSVYFVGKTPNIEPQGYSE-----EEEEEEDV 113
DDB 59 knphigfdlfdkfelshstkttsvftgkykv-----egfcedemdlstdedeelnv 114
QY 114 PA-----GNA-----AKAVAKP-KAKPAEVKPAVDDDEDESDDSGMDEDDSD----- 154
DDB 115 pvvkengkadgkqkqsekavapskspdkskdaddsdededetddsddegl 174
QY 155 -----CEDSEEEPTPKPKPASSKKRANETTP-KAPVSAKAKAVAVTP--Q 196
DDB 175 sseegddsddeedtptpkpvgkakraessvltkplsdkkakva-tpssq 233
QY 197 KTDEKK-----KGGKAANQSPKASQVSCGCKKTFNSGNALSH 235
DDB 234 ktggkgaavhvathpapakgtivnndksvskpksapkgsgvpcskscsksfiseta 293
QY 236 HNAKAKH 241
DDB 294 hskakh 299

RESULT 10
AAY28802
ID AAY28802 standard; protein; 305 AA.
XX AC AAY28802;
XX 13-JAN-2000 (first entry)
DE Maize histone deacetylase-6.
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
OS Zea mays.
XX WO9951731-A2.
PN 14-OCT-1999.
XX 02-APR-1999; 99WO-US07370.
XX 03-APR-1998; 98US-0080563.
XX (PION-) PIONEER HI-BRED INT INC.
XX Baldwin DA, Briggs SP, Crane VC;
XX WPI; 1999-611038/52.
XX N-PSDB; AAX90842.
XX New deacetylase genes, used for producing transgenic plants which have
XX increased disease resistance -
XX Claim 1; Page 72-73; 87pp; English.
XX The present sequence is maize histone deacetylase encoded by HD cDNA
XX belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
XX modifications, may be localised to promoters targeted for repression by
XX other proteins that associate with HD and specifically bind regulatory
XX elements in promoter DNA. The HD nucleotide sequence can be used for
XX producing transgenic plants with increased disease resistance.
XX Additionally, compositions find use in screening for toxins that affect
XX pathogenicity and in determining which disease response promoters are
XX regulated by histone deacetylase.
SQ Sequence 305 AA;

Query Match 38.1%; Score 487; DB 20; Length 305;
Best Local Similarity 41.2%; Pred. No. 8e-36;

```


PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 24-JUN-1999; 99US-0140823.
 PR 28-JUN-1999; 99US-0140991.
 PR 29-JUN-1999; 99US-0141287.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151330.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.7%; Score 405; DB 21; Length 126;
Best Local Similarity 71.7%; Pred. No. 6.5e-29;
Matches 81; Conservative 11; Mismatches 19; Indels 2; Gaps 2;

OY 1 MFTWGVKSGKPVTPTEGILIHVSQSLGCKNKKGEFVPLHVKNQNLVGLTST 60
|||||
Db 1 mefWGVKSGKpvtPteegILIHVSqslgckNkkgefVplhVknqnlvGlTst 60
|||||

OY 61 ENIPQPCDLVDFDKFELSHTWCKGSVY-FVGVKTPNIEPQGVSEEEEEE 112
|||||
Db 61 enipqpcdlvdfkelsf-tlgekevfilldklptishkallrkrkrk 112
|||||

RESULT 13
AAG25823
ID AAG25823 standard; Protein; 208 AA.
XX AC AAG25823;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30043.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0310439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.3%; Score 374.5; DB 21; Length 208;

Best Local Similarity 39.4%; Pred No. 7e-26; Mismatches 29; Conservative 86; Indels 41; Gaps 5;

QY 1 MEFWGIEVSKGKPVTVTPEEGILIHVSQASLGCKKNGKEFVPLHVKNQNLVLTST 60

Db 1 meifgvevknkplhldpgldrlvhisqvalgesknvteplqlyvtvgsdklligtish 60

QY 61 ENIPQLFCDLVFDKREFELSHTWKSGSVYFVGKTPNIEVQGYSEEEEEEEVPACNAAK 120

Db 61 ekfpqlsteivlernfalshwtknsgsvffsgykvdsdp-----epedliddqleaa-gfk 115

QY 121 AVAKPKAKPA-----EVKPAVDDEEDSDSGMDEDDSDS----- 153

Db 116 aapksaakqvnfqlpnedvkakqddadgdseddssdddsensgdeeeekvtaesdseed 175

QY 154 ----DGEDSEEEPTPKPKASSKKRANETTPKAPVSAK 187

Db 176 dsdsdeeddsseetpkpkpeepkksae-----pnssk 208

RESULT 14

AAG09171

ID AAG09171 standard; Protein; 203 AA.

AC AAG09171;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 6998.

DE Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

```
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 23.9%; Score 306; DB 21; Length 203;
Best Local Similarity 36.8%; Pred. No. 1e-19;
Matches 86; Conservative 28; Mismatches 80; Indels 40; Gaps 12;

QY 1 MEWGVIEVSGKPVTVTPERGIIHVHSQASLGEC-KNKKGEFVPLHVKNVGN--QNVLVLT 57
Db 1 metwgieikpgkpfkvgkfmgfmgvhasqvtlgdvkvvkddetfavvykkgiddengfmign 60
QY 58 LSTENIPOLCDLVFOKEFELSHTWCKGVSFYFYGYKTPNIEPOGYSEEEEEEEVYPAGN 117
Db 61 ls-qkfpqfidlylghfeishn-stsvylligrt--fdafdeldeidselideym 116
QY 118 AAKAVAKPAKPAEVAKPAVDDEEDSD--GMDEDD--SDGDSDEEPTPKKPASSKK 173
Db 117 eqqlaalpq--neinp---eeddesdsdemgldedddssdeedveaeaplkvappskk- 169
QY 174 RANETPKAPVSAKAKAVATPQKTEKKGGKAAQNS-PKASQVSCG-SCKK 225
Db 170 -----mpngafeiakggknkssggkkrctpcgpcskk 203

RESULT 15
AAG17704
ID AAG17704 standard; Protein; 155 AA.
AC AAG17704;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 18826.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18826.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
```

```
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

Query Match 20.0%; Score 255.5; DB 21; Length 155;
Best Local Similarity 47.3%; Pred. No. 2.6e-15;
Matches 71; Conservative 10; Mismatches 32; Indels 37; Gaps 9;
QY 89 FVGKTPNIEFGY-SEEEEEEEVPA-----GNAAKAV-----AKPKAKPAEVKP 134
Db 6 faskkkpnieqddflssddedvpeavpapaplatvLangnaqaavvkadskpkakpaevkp 65
QY 135 AVD-----DEEDESDD-----CM--DEDDSDG---EDS---EEEETPKKPAS-S 171
Db 66 aeekpesdecdedddeseeeddskgmvdvddsddeeddeeddeeeetpnkpepin 125
QY 172 KKRANETTPKAPVSAKAKVAVTPQKTDEK 201
Db 126 kkrpnesvxtipvsgkkakpaapastpqk 155

Search completed: May 2, 2002, 22:03:52
Job time: 6540 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:26:44 ; Search time 60.74 Seconds
(without alignments)
98.523 Million cell u

Title: US-09-645-337-6
 Perfect score: 1279
 Sequence: 1 MEFWGIEVKSGKPVTVTEE.....TFNSGNALESHNKAHAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCUTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	513.5	40.1	311	4	US-09-282-305-16	Sequence 16, Appl
2	499	39.0	302	4	US-09-282-305-14	Sequence 14, Appl
3	487	38.1	305	4	US-09-282-305-12	Sequence 12, Appl
4	446	34.9	285	4	US-09-282-305-18	Sequence 18, Appl
5	224	17.5	412	2	US-08-741-134-2	Sequence 2, Appl
6	166.5	13.0	714	2	US-08-990-114-3	Sequence 3, Appl
7	166.5	13.0	714	4	US-09-241-333-3	Sequence 3, Appl
8	149	11.6	257	4	US-09-538-871-4	Sequence 4, Appl
9	139.5	10.9	411	2	US-08-741-134-6	Sequence 6, Appl
10	136.5	10.7	292	4	US-09-538-871-2	Sequence 2, Appl
11	129.5	10.1	1085	1	US-08-431-080-28	Sequence 28, Appl
12	129.5	10.1	1085	2	US-08-938-534-28	Sequence 28, Appl
13	127	9.9	565	4	US-08-961-083-218	Sequence 218, Appl
14	126.5	9.9	258	4	US-08-961-083-90	Sequence 90, Appl
15	126.5	9.9	723	1	US-07-814-964-11	Sequence 11, Appl
16	126.5	9.9	723	1	US-08-258-442-11	Sequence 11, Appl
17	126.5	9.9	723	1	US-08-328-809-6	Sequence 6, Appl
18	126.5	9.9	723	5	PCR-US92-11107-11	Sequence 11, Appl
19	121.5	9.5	240	2	US-08-114-558A-8	Sequence 8, Appl
20	121.5	9.5	240	3	US-08-559-397A-14	Sequence 14, Appl
21	121.5	9.5	251	2	US-08-766-738-3	Sequence 3, Appl
22	121	9.5	542	1	US-08-701-380-2	Sequence 2, Appl
23	121	9.5	542	3	US-09-032-365A-13	Sequence 13, Appl
24	117.5	9.2	226	1	US-08-431-080-26	Sequence 26, Appl
25	117.5	9.2	226	2	US-08-938-534-26	Sequence 26, Appl
26	116.5	9.1	739	4	US-09-022-983-2	Sequence 2, Appl
27	115	9.0	764	1	US-08-375-300-4	Sequence 4, Appl

RESULT 2

RESOLUTION
US-09-282-305-14

28	115	9.0	764	3	US-09-177-431-4	Sequence 4, Appl
29	115	9.0	764	5	PCR-US95-16930-4	Sequence 4, Appl
30	115	9.0	1089	1	US-08-375-300-2	Sequence 2, Appl
31	115	9.0	1089	3	US-09-177-431-2	Sequence 2, Appl
32	115	9.0	1089	5	PCR-US95-16930-2	Sequence 2, Appl
33	113	8.8	674	3	US-08-893-852A-1	Sequence 1, Appl
34	110.5	8.6	740	4	US-09-022-983-5	Sequence 5, Appl
35	109.5	8.6	160	2	US-08-726-306A-68	Sequence 68, Appl
36	109.5	8.6	251	2	US-08-766-738-1	Sequence 1, Appl
37	109.5	8.6	1187	1	US-08-320-559-28	Sequence 28, Appl
38	109.5	8.6	1187	3	US-08-545-860D-28	Sequence 28, Appl
39	109.5	8.6	1187	5	PCR-US94-04496-28	Sequence 26, Appl
40	109.5	8.6	1210	1	US-08-320-559-26	Sequence 26, Appl
41	109.5	8.6	1210	3	US-08-545-860D-26	Sequence 26, Appl
42	109.5	8.6	1210	5	PCR-US94-04496-26	Sequence 26, Appl
43	109.5	8.6	1706	2	US-08-459-568-2	Sequence 2, Appl
44	109.5	8.6	1706	3	US-08-399-411-2	Sequence 2, Appl
45	109.5	8.6	1706	5	US-08-516-859A-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT      1
US-09-282-305-16
; Sequence 16, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Zea mays
US-09-282-305-16

```

[illegible]

```

; Sequence 14, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/22
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,5
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 302
; TYPE: prt
; ORGANISM: Zea mays
; US-09-282-305-14

```

Query Match 39.08; Score 499; DB 4; Length 302;
Best Local Similarity 40.58; Pred. NO. 1.7e-38;
Matches 124; Conservative 36; Mismatches 74; Indels 72; Gaps 11;

[illegible]

RESULT 3

```

US-09-282-305-12
? Patent 12, Application US/09282305
? Sequence No. 6287843
? GENERAL INFORMATION:
? APPLICANT: Baldwin, Donald A.
? APPLICANT: Briggs, Steven P.
? APPLICANT: Crane, Virginia C.
? TITLE OF INVENTION: Male Histone Deacetylase
? FILE REFERENCE: 5718-44.
? CURRENT APPLICATION NUMBER: US/09/282,305
? CURRENT FILING DATE: 1999-03-31
? PRIOR APPLICATION NUMBER: 60/080,563
? PRIOR FILING DATE: 1998-04-03
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 12
? LENGTH: 305
? TYPE: PRT
? ORGANISM: zea mays
US-09-282-305-12

```

Query Match 38.1%; Score 487; DB 4; Length 305;
Best Local Similarity 41.2%; Pred. NO. 2.2e-37;
Matches 127; Conservative 37; Mismatches 74; Indels

QY	1	MEFWGIEVKGKGVYVTPREGILLHYVSOASLGCKNKKGEFVPLHVKVGKGNONI.VI.GT.LST 60
Db		: : : : : :
QY	1	MEFWGLVPGSTVKCEPGYGVFLHLSQAALGE--SKKSDNALMWYKIDDKLAIGT.LSV 58
Db		: : : : : :
QY	61	ENIPOLPCDLVFOKEPELSHTWKGSGSVYFVGYKTPNIEPOCYSE-----EEEEEEVV 113
Db		: : : : : :
QY	59	DKNPHIQFDLIFDKPEPELSHTKTTSVFFTCYKV---EQPFEDEMDLSDSEDFELNV 114
Db		: : : : : :
QY	114	PA-----GNA-----AKAVXP-KAKPAEVPAYVDEDEDESDSGMDEDDSD----- 154
Db		
QY	115	PVVKENGKADKKKOKSGFKAAVAPSKSSPSKSKDDDDSDDEDTDDSDDETDSDDEGL 174
Db		: : : : : :
QY	155	-----GDSFSEERPTPKKPASSKKRANETTP-KAPYSAKKAKVATP-Q 196
Db		: : : : :
QY	175	SPREGDDSDSDDEDTSDDEEDTPTPKKEPVGKKRAEVSVLKTPUSDKKAKVA-TPSSQ 233
Db		: : : : : :
QY	197	KTDPKK-----KGGKAAN-----QSPKSA--SOVSCSCCKKTTFSGNALF.SHN 237
Db		: : : : : :
QY	234	KTGCKKCAAVHVA TPHPKCKTIVNNDKSVKPSAPKSGVPCKSCSKSP.ISETAI.QAHS 293
Db		: : : : : :
QY	238	KAKHAAK 245
Db		: : : : : :
QY	294	KANNGASE 301
Db		: : : : : :

RESULT 4

```

US-09-282-305-18
: Sequence 18, Application US/09282305
: Patent No. 6287843
: GENERAL INFORMATION:
: APPLICANT: Baldwin, Donald A.
: APPLICANT: Briggs, Steven P.
: APPLICANT: Crane, Virginia C.
: TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
: FILE REFERENCE: 5718-44,
: CURRENT APPLICATION NUMBER: US/09/282,305
: CURRENT FILING DATE: 1999-03-31
: PRIOR APPLICATION NUMBER: 60/080,563
: PRIOR FILING DATE: 1998-04-03
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 18
: LENGTH: 285
: TYPE: PRT
: ORGANISM: Zea mays
US-09-282-305-18

```

Query Match 34.9%; Score 446; DB 4; length 285;
Best Local Similarity 41.9%; Pred. No. 1.2e-33;

	Matches	122:	Conservative	31:	Mismatches	84:	Indels	54:	Gaps
Qy	1	MEFWGIEVKGKPVVTPEKGIILLHVSQASIGECKNKGFEVPLHVKGVGNONI.VLGLTST	60						
Db	1	MEFWGEEVKGATVSCVKVGDCI.VIHLISQAALGGPK-KASENAIVSVKIDDKKILVGLTSLV	59						
Qy	61	ENIPOLFCDLAFDEKEFELSHTWCKGGSVYFVGYKTP----	NIEPQGYSEEEEEEEEPAG	116					
Db	60	EKHPIQISCDLIVFDKDFELSHNSKTASVFCGYKSPVPLFHSDEGHSUDEVEFDPDILPMQ	119						
Qy	117	N-----AAKAVAK----	PKAPAEVKPAVDDEDESDGMDDEDDSDGDP-	157					
Db	120	NNEIKISTAKVPVKGVIQNADEDETSSGDDDDFTTDSSENSEDESDSDSDTSDTSD	179						
Qy	158	---SEEEPTPKP-ASSKKRANET--TPKAPVSAAKKAYVTPKTDK-----	KKG	204					
Db	180	DSGSEEQTPPKTEVVVGKKRAIEAETP-----SGKKAKSKQAQKTDGVKYSTSHPAKQS	235						
Qy	205	GK-----AANQSPKASASOVSCGCKTKTFNSGNALSHENKA-KHAA	243						

TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-741-134-6

Query Match 10.9%; Score 139.5; DB 2; Length 411;
Best Local Similarity 27.1%; Pred. No. 4.2e-05;
Matches 60; Conservative 22; Mismatches 98; Indels 41; Gaps 8;
Qy 51 QNLVLGTLSTENIPQLFCDLVFKEFELSH-TWKGSGVYFVG--YKTPNIEPOGYSEEE 107
Db 118 QESVLLTSPAEAOQOQLDLTITPEEVQFIVTGSVAISLGNVVKHPFTPMGVEGEDE 177
Qy 108 EEEVEVPAGNAKAVAKPAEVPKPAVDDEDE-----SDSGMDDESDGSDSEEE- 162
Db 178 DEDADIYDSEDYDLTPDEDEIIGDMDDDEEEVRIEEVQDEDEEDNDGEBQEEEE 237
Qy 163 -----PTPKPKASSKKRANETTPKAPVSAKKAHVAVT-----PKTDEKKKGKA 207
Db 238 EEEQKEVPEPKSKKKEKKHKEEKEEKKAKKVKVEFKDLEEGTPPKSKKQDK- 296
Qy 208 ANQSPKASQVSCGCKTKTNSGALSHN-----KAKHAA 243
Db 297 --HKPKS-----KVEGGIVIEDRTIGDGPQAKRGA 325

RESULT 10
US-09-538-871-2
Sequence 2, Application US/09538871
Patent No. 6358718
GENERAL INFORMATION:
APPLICANT: Olson, Mark
APPLICANT: Szebenl, Atilla
TITLE OF INVENTION: METHODS FOR STABILIZATION AND
TITLE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS B23
FILE REFERENCE: 13121.000202
CURRENT APPLICATION NUMBER: US/09/538,871
CURRENT FILING DATE: 2000-03-30
EARLIER APPLICATION NUMBER: 60/126,910
EARLIER FILING DATE: 1999-03-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 292
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of artificial sequence:/No. 6358718e -
US-09-538-871-2

Query Match 10.7%; Score 136.5; DB 4; Length 292;
Best Local Similarity 25.4%; Pred. No. 5e-05;
Matches 52; Conservative 34; Mismatches 90; Indels 29; Gaps 8;
Qy 24 IHVQSALGECNKKKGFEVPLHVYKGNQNLVLTSTENIPQLFCDLVFKEFELSH--- 80
Db 57 LHIVEA---EAMNTEGS--PIKV-----TLATLKMSVQPTVSVHL---GGFEITPPVV 101
Qy 81 ---TWKGSGVYFVGKTPNIEPOGYSEEEVEEVEVPAGNAKAVAKPAEVPKPAVD 137
Db 102 LRLKCGSGPI---SGHVLVAVEADESEDEDEEDVKLLGMSKRSAPGGNPKQKVKYLD 159
Qy 138 DEEDFSDGMDDEDDSDGSESEEEPTPKPASSKKRANET-----TPKAPVSAKKAHVAV 193
Db 160 EDDDEDEDEDEDEDEDEDETEKVPV--KKSVRTPAKNAOKSNQNGKDLKPST 217

Qy 194 TPQKTDEKKKGKKAANQSPKASQV 218
Db 218 PRSKQSEFKQEKTPKTPKGPSSV 242
RESULT 11
US-08-431-080-28
Sequence 28, Application US/08431080
Patent No. 5698686
GENERAL INFORMATION:
APPLICANT: Gottschling, Daniel E.
APPLICANT: Singer, Miriam S.
TITLE OF INVENTION: Telomerase Compositions and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/431,080
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SN 08/326,781
FILING DATE: October 20, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:155/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-431-080-28

Query Match 10.1%; Score 129.5; DB 1; Length 1085;
Best Local Similarity 23.5%; Pred. No. 0.0013;
Matches 62; Conservative 29; Mismatches 96; Indels 77; Gaps 9;
Qy 15 TVTPEEGILIHVSQSLGEC-----KNKKGFEVPLHVYKGNQNLVLTSTENIPQLFCDL 70
Db 47 TTRPRRFLSIYSESSLSVSDSDKNKSTN--PHKIKRKAANI-----SNNQCKKSK 97
Qy 71 VFDEKELSHITWKGSGVYFVGKTPNIEPOGYSEEEVEEVEVPAGNAKAVAKPAKPA 130
Db 98 LIQRQINDNDEGTSSDYQA--VTDGESENEEEEEEEED-----137
Qy 131 EVKPAVDDEDESDSGMDDEDDSDGSESEEEPTP--KKPASSKKA-----NE 177
Db 138 -----DDEDDDDDDGSD 191
Qy 178 TTPKAPVSAKKAHVAVTPQKTDEKKKGKKAANQSPKASQVSCGCKTKFNS----- 229
Db 192 LYSRSNKNKSVKLSPPKNEEEQ--KKEKEKEEEOQKQOUSNKEVNGSGTTTQQ 249
Qy 230 -----GNALESHNK 238
||| | | |

[illegible]

```

US-08-961-083-90
: Sequence 90, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.

```

```

; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 258 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-90

```

```

Query Match          9.9%; Score 126.5; DB 4; Length 258;
Best Local Similarity 30.5%; Pred. No. 0.00036;
Matches 40; Conservative 17; Mismatches 57; Indels 17; Gaps 4;

QY 98 EPQGYSEEEEEEEVPAGNAAKAVAKPKAKP-AEVKPAVD-----DEEDSDSDGMD 149
Db 9 QEGARTNQVTEEAAPKEEAPKTESEKPEKSEKPTDDTLPKVEEGKDSAEPAPE 69

QY 150 EDDSDGSDSEEEPTPK-----KPASSKKRANETTPKAPVSAKKAKVAVTPKTDK 202
Db 69 EVGGEVSEKPEKAVKVPESQPSDKPAESKVEQAGEVPAPREDEKA--PVEPEKPEAP 126

QY 203 KGGKAANQSPK 213
Db 127 EEEKAVEETPK 137

```

```

RESULT 15
US-07-814-964-11
; Sequence 11, Application US/07814964
; Patent No. 5359047
; GENERAL INFORMATION:
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pili, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellelt, Patti
; APPLICANT: Essigmann, John M.
; APPLICANT: Lippard, Stephen J.
; TITLE OF INVENTION: DNA Structure Specific Recognition
; TITLE OF INVENTION: Protein and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA

```

```

; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,964
; FILING DATE: 19911226
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539,906
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4787AAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 11
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; IMMEDIATE SOURCE:
; CLONE: Drosophila SSRP (predicted)
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 458..507
; OTHER INFORMATION: /label= Acidic
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 518..547
; OTHER INFORMATION: /label= Basic 1
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 547..620
; OTHER INFORMATION: /label= HMG-box
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label= Basic 11
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label= Mixed Charge
; US-07-814-964-11

```

```

Query Match          9.9%; Score 126.5; DB 1; Length 723;
Best Local Similarity 26.4%; Pred. No. 0.0014;
Matches 58; Conservative 36; Mismatches 85; Indels 41; Gaps 12;

QY 54 VLGTLTSTENIPQLFCDLVFDKFEFSLHTWKGSVFVGVYKTPNI-----EPQGY---- 102
Db 402 IFSSIEKEEYAKLF-DYITQKKLHYSN-MGDK---SGYKDVDFGSDNENEPDAYLARL 456

QY 103 SEEEEEEVEVPAGNAAKAVAKPKAPVAKPAVDEED---ESDSDGMD-----DS 153
Db 457 KAAEREKEEDDDDDSDERSTDEDFKPNEDSDVAEEYDSNVESDDO--DSDASGGGDS 515

QY 154 DGEDSEEEPTPKPASSKKRANETTPKAPVSAKKAKVAVTPKTD-----DEKKKG 205
Db 516 DGAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 573

QY 206 KANQSPKSAVSOVSCGCKKTFNSGNALSHNKAHAKHAAK 245
Db 574 KRENPKIK-VTEIA----KKGEMWKLKDKSKWEDAAAK 608

```

Search completed: May 2, 2002, 22:08:41
Job time: 6117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 20:16:12 ; Search time 205.36 Seconds

(without alignments)
114,637 Million cell updates/sec

Title: US-09-645-337-6

Perfect score: 1279

Sequence: 1 MEFWGIEVKSGRPVTVPEE.....TFNSGNALSHNKAHMAAK 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-71:**

1: p1r1:**

2: p1r2:**

3: p1r3:**

4: p1r4:**

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	594	46.4	305	2 T52287	probable histone d
2	498.5	39.0	296	2 T48401	histone deacetylase
3	482	37.7	307	2 T04141	histone deacetylase
4	224	17.5	412	2 A55320	immunophilin FKBP4
5	188.5	14.7	357	2 JC4090	FK506-binding 39k
6	170.5	13.3	299	2 A29681	nucleophosmin - At
7	170	13.3	694	1 DNGHML	nucleolin - chicken
8	169	13.2	294	1 DNGHML	nucleophosmin - ch
9	166.5	13.0	713	2 A27441	nucleolin - Chinese
10	165.5	12.9	990	2 T51618	nucleolar phosphop
11	164	12.8	487	2 T10215	hypothetical prote
12	157	12.3	1110	2 T51116	NF-180 - sea lamp
13	155	12.1	294	1 A32915	nucleolin homin - hu
14	153.5	12.0	635	2 T09648	nucleolin homolog
15	153.5	12.0	707	1 DMS	nucleolin - mouse
16	153	12.0	707	2 A35804	nucleolin - human
17	151.5	11.8	712	2 JH0148	nucleolin - rat
18	150.5	11.7	281	2 JC4295	heat-shock protein
19	149.5	11.6	644	2 S55395	neurofilament prot
20	149	11.6	257	2 A34168	nucleolar phosphop
21	149	11.6	292	2 A28939	nucleophosmin - ra
22	148	11.6	296	2 A41730	nucleophosmin NO38
23	146.5	11.5	798	2 T33022	hypothetical prote
24	145.5	11.4	734	2 B42680	nucleolus-cytoplas
25	144	11.3	292	2 T52858	hypothetical prote
26	140.5	11.0	649	2 D85135	neurofilament medi
27	140.5	11.0	798	2 T50479	zinc metalloprotei
28	140.5	11.0	1881	2 H95076	peptidylprolyl iso
29	139.5	10.9	411	2 S48647	

30	138.5	10.8	971	2 T19431	hypothetical prote
31	138	10.8	699	2 T38073	nucleolar phosphop
32	138	10.8	705	2 S32644	nucleolin - Africa
33	137	10.7	590	2 S63193	hypothetical prote
34	136.5	10.7	292	2 A36089	nucleolar phosphop
35	136.5	10.7	611	2 T06458	nucleolin homolog
36	136.5	10.7	723	2 A48217	single-strand DNA/
37	136	10.6	651	2 S18874	nucleolin - Africa
38	135.5	10.6	557	2 A96527	probable num1 prot
39	134.5	10.5	280	2 A61047	ectodermal (ect) -
40	134.5	10.5	325	2 T18283	hypothetical prote
41	134	10.5	1320	2 UC5630	TCOF1 protein - mo
42	133.5	10.4	1876	2 E97944	zinc metalloprotei
43	133	10.4	532	2 T06029	hypothetical prote
44	132.5	10.4	971	2 T24866	hypothetical prote
45	132	10.3	374	2 C88734	protein P32E10.6

ALIGNMENTS

RESULT 1
T52287
Probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #extl_change 20-Oct-2000
C:Accession: T52287
R:Danf, M.; Haas, H.; Ioldi, P.
Submitted to the EMBL Data Library, January 1998
A:Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2
A:Reference number: Z26016
A:Accession: T52287
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1305 <DAN>
A:Cross-references: EMBL:AF044914; PIDN:AAC02339.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: HD2
C:Keywords: hydrolase

Query Match 46.4%; Score 594; DB 2; Length 305;
Best Local Similarity 50.7%; Pred. No. 6.5e-11;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;
QY 1 MEFWGIEVKSGRPVTVPEEGILLIHSQASLGCKKNGKGFVPIHYKVGQNIIVICTLST 60
DB 1 MEFWGAVVTEPKNAIKVTPEDSLVHSQASL-DCTVKSSESVALSTVGGAKLVIGTLSS 59
QY 61 ENIPQLCDLVDPKEPELSHTWCKGSYFYGYKTPNIEPGCY-SSEEEEEEVPA---- 115
DB 60 DKFPQISFDLVFDEKPELSHSGTANVHPTGYKSPNIBDDPISDDDEVPVAPAPPT 119
QY 116 -----GNAKAV-----AKPKAPAEVPAVD-----DEEDSDSD-----GM--D 149
DB 120 AVTANGAAGAAVYKADIKPKAPAEVPAEKEPESDESDDESDDESDDESEKMDVD 179
QY 150 EDDSDG---EDS---EEEPPTPKPAS-SKKRANETTPKAPYSAKKAKVAV----TQKTD 199
DB 180 EDDSDDEEDSDSEDEEEETPKRREPINKRPNESVSKTSVSKKKKPPAAPAAPSTQKTE 239
QY 200 EK-----KKGKA---ANOSPKSAOYSCG-SCKKTFNSGNAL-ESHNNKAKH 241
DB 240 KKKGGHATPHPAKKGKSPYMANQSPKSGOSGCGNNKKPPNSCKGKFGGNNKGSN 297
RESULT 2
T48401
Histone deacetylase-like protein - Arabidopsis thaliana
N:Alternate names: protein F17C15.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #extl_change 20-Apr-2000

A:Title: FKBP39, a Drosophila member of a family of proteins that bind the immunosuppressant
 A:Reference number: J04090; MUID:95278752
 A:Accession: J04090
 A:Molecule type: mRNA
 A:Residues: 1-357 <THE>
 A:Cross-references: EMBL:Z46894; NID:9600423; PIDN:CAA86996.1; PID:9600424
 A:Gene: fkbp39
 A:Cross-references: FlyBase:FBgn0013269
 C:Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomerase
 F:89-99/Region: acidic
 F:252-357/Region: FK306 binding *status predicted
 F:269-316/Domain: BKBP-type peptidylprolyl isomerase homology <PI>

Query Match 14.7%; Score 188.5; DB 2; Length 357;
 Best Local Similarity 25.1%; Pred. No. 4.6e-05;
 Matches 71; Conservative 43; Mismatches 110; Indels 59; Gaps 10;

OY 3 FWGIEVSKGKPYVTPPEGIILHVSQASLGECNKKCFEPLHVGQNLVLGTLSTEN 62
 DB 4 FVGUMKKERKRSQITLIS--FHISGVAL-----DKGQKATLYLAKEQETIVATV-TKA 55
 OY 63 IPQLFCDLVFDK-EPELSHTWCKGSGYVFGYKTPNIEPOGSEEESEEEVPAAGNAKA 121
 DB 56 IPOVALDINFSKGRIMFYTAGDASVSLGYL-----HDIDSEDEDDDDMTIENLINS 109
 OY 122 VAKPKAKPA-----EVKPAVDEED-----ESDSGDMD----- 149
 DB 110 KAIKMSKSEDEDEDENEGEDEEDTDDSQIIEYESFLENGEDEDDDDDEDENESEGE 169
 OY 150 EDDSDGDESEEEPTPKRP-----ASSKRANETTPKAPYSAKAKAVATVPOKTDKK 202
 DB 170 EDEQSDSDSEAEEDQPKRKRVAKLSPGASAKKSGKQNGVA-----KKREAAQQQKKKKRP 224
 OY 203 KGGKAAKNSPKSASQVSCGCKKTNSGNALESHKAKAKAAK 245
 DB 225 EAKKQPKAPKAPKAKQAPSKDPRTITGGVKTVDQVVGKGEAK 267

RESULT 6
 A29681
 A:Title: Xenopus laevis (African clawed frog) nucleophosmin - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 13-Aug-1999
 A:Accession: A29681
 R:Schmidt-Zachmann, M.S.; Huegler-Doerr, B.; Franke, W.W.
 EMBL J. 6, 1881-1890, 1987
 A:Title: A constitutive nucleolar protein identified as a member of the nucleoplasmic
 A:Reference number: A29681; MUID:88004394
 A:Accession: A29681
 A:Molecule type: mRNA
 A:Residues: 1-299 <SCH>
 A:Cross-references: GB:X05496; NID:964924; PIDN:CAA29046.1; PID:964925
 C:Superfamily: nucleophosmin
 C:Keywords: molecular chaperone; nucleus

Query Match 13.3%; Score 170.5; DB 2; Length 299;
 Best Local Similarity 31.5%; Pred. No. 0.00053;
 Matches 56; Conservative 20; Mismatches 67; Indels 35; Gaps 7;

OY 83 GKGSYFVGKTPNIEPGYSEEESEEEVPAAGNAKAVAKPAKPAVEVPAVDEDE 142
 DB 107 GSGPYVSSQHLVALEDDSEDEDEHEHPSPK-NAKRIAPDSAKYVRKTRLEEEED 165
 OY 143 SD-SQGMDESDSGDSEEEPTP-KKPAASSKRA-----NETTPKAPVS 185
 DB 166 SDEDDDEDDDEDEDEDEETPVKTKTSTKSKAAQKLNHGKASALSTQKTPKTDQ 225
 OY 186 AKK-----AAVATPOKTDKK-----KGGKAAKNSPKSASQVSCGCKKTEN 228
 DB 226 KKKQTKPTPTKPTPLSSSEIKAKMOTYLEKGNVLPKVEKPAVYK--NCFRTEN 281

RESULT 7

DNCNLT
 N:Alternate names: nucleolar protein C23
 C:Species: Gallus gallus (chicken)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
 A:Accession: S08414; S10766; A32725; I50397; B30099
 R:Maridor, G.; Nigq, E.A.
 Nucleic Acids Res. 18, 1286, 1990

A:Title: cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar pr
 A:Reference number: S08414; MUID:90206792
 A:Accession: S08414

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-694 <MAR>

A:Cross-references: EMBL:X17199; NID:963710; PIDN:CAA35060.1; PID:963711

R:Maridor, G.; Krek, W.; Nigq, E.A.
 Biochim. Biophys. Acta 1049, 126-133, 1990

A:Title: Structure and developmental expression of chicken nucleolin and NO38: coordi
 A:Reference number: S10766; MUID:90304215

A:Accession: S10766

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-694 <MAR2>

R:Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigq, E.A.
 Cell 60, 791-801, 1990

A:Title: Identification of major nucleolar proteins as candidate mitotic substrates o
 A:Reference number: A32725; MUID:90182668

A:Accession: A32725

A:Molecule type: protein

A:Residues: 56-62; 63-109; 197-214 <PET>

R:Boer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigq, E.A.
 Cell 56, 379-390, 1989

A:Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.

A:Reference number: I50397; MUID:89119560

A:Accession: I50397

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 407-418, 'T', 521-694 <BOR>

A:Cross-references: GB:M21791; NID:9212411; PIDN:AAA48983.1; PID:9212412

C:Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorg
 C:Superfamily: nucleolin; ribonucleoprotein repeat homology

C:Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding

F:1-247/Domain: acidic <AC2>

F:54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])

F:254-262/Region: nuclear location signal

F:267-275/Region: nuclear location signal

F:282-347/Domain: ribonucleoprotein repeat homology <RRM1>

F:372-435/Domain: RNA binding *status predicted <RNA1>

F:373-448/Domain: ribonucleoprotein repeat homology <RRM2>

F:463-525/Domain: RNA binding *status predicted <RNA2>

F:463-538/Domain: RNA binding *status predicted <RRM3>

F:554-618/Domain: RNA binding *status predicted <RNA3>

F:555-631/Domain: ribonucleoprotein repeat homology <RRM4>

F:555-631/Domain: RNA binding *status predicted <RNA4>

F:632-694/Domain: glycine/arginine-rich <GR>

F:56,63,70,77,85/binding site: phosphate (Thr) (covalent) (by cdc2 kinase) *status pr

Query Match 13.3%; Score 170; DB 1; Length 694;
 Best Local Similarity 33.9%; Pred. No. 0.0013;
 Matches 60; Conservative 16; Mismatches 61; Indels 40; Gaps 8;

OY 84 KGSYFVGKTPNIEPOGSEEESEEE-----EEVPAAGNAKAVAKP--KAK 128
 DB 151 KKSAAVPAKPAVPAKQSEEESEEEDEDEDEDEDEDEADMTTPAVKPKTPAKAT 210
 OY 129 PAVKPAVDEDESDSGMDESDSGDESDSEEPTRKPRAS-----SKKRANETPKA 182
 DB 211 PAKAKAEESEDEDEDEDEDEDEDEDEDESEDEKPVKPAKPKKEMAKSAPEA 269

A:Accession: B26080
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 'P', 214-226 <CHAI>
 A:Cross-references: GB:M31004; GB:M12775; NID:g190239
 A:Experimental source: clone hpb2
 A:Accession: A26080
 A:Molecule type: mRNA
 A:Residues: 'P', 214-215, 'S', 217-218, 'SSS', 222-294 <CHAI>
 A:Cross-references: GB:J02590; NID:g190237; PIDN:AAA36473.1; PID:g190238
 C:Genetics:
 A:Gene: GDB:NM1
 A:Cross-references: GDB:138455; OMIM:164040
 A:Map position: 5q35-5q35
 C:Superfamily: nucleophosmin
 C:Keywords: nucleolus; nucleus: phosphoprotein
 F:120-132/Region: glutamic acid-rich
 F:152-158/Region: nuclear location signal
 F:161-188/Region: aspartic acid/glutamic acid-rich
 F:191-198/Region: nuclear location signal
 F:125/Binding site: phosphate (Ser) (covalent) *status predicted

Query Match 12.1%; Score 155; DB 1; Length 294;
 Best Local Similarity 28.3%; Pred. No. 0.0051;
 Matches 43; Conservative 25; Mismatches 56; Indels 28; Gaps 5;

QY 83 GKSVYFVGVKTNIEPGYSEEEEEE-----EEVPAQNAKAVAKPAKPAEYK 133
 DB 105 GSGPVHISQHLAVAEDEASEDEDEDYKLSISCKRSPGCG-----SKYQKVKVLA 159
 QY 134 PAVDDEEDSDSCMDSDSDGSEDEEPTPK-----PASSKRAETTPKAVSAK 187
 DB 160 ADEDDDDDEDEDDDDDDDDDEDEAEKAPYKSIKRTPAKNAKSNQ-----NGK 212
 QY 188 KAKVAVTPQ-KTDEKKKGGCAANSPKASQV 218
 DB 213 DSKPSTPKSGQSFRRKQKPTKTPKPSV 244

RESULT 14
 T09648
 nucleolin homolog nm1 - alfalfa
 N:Alternate names: nm1 protein
 C:Species: Medicago sativa (alfalfa)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T09648
 R:Boegre, L.; Jonak, C.; Mink, M.; Meskene, I.; Traas, J.; Ha, D.T.C.; Swoboda, I.; Plant Cell 8, 417-428, 1996
 A:Title: Developmental and cell cycle regulation of Alfalfa nm1 a plant homolog of th
 A:Reference number: Z16796; MUID:96361876
 A:Accession: T09648
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-635 <R0>
 A:Cross-references: EMBL:X88845; NID:g1279562; PIDN:CAA61298.1; PID:g1279563
 C:Genetics:
 A:Gene: nm1
 C:Superfamily: nucleolin; ribonucleoprotein repeat homology

Query Match 12.0%; Score 153.5; DB 2; Length 635;
 Best Local Similarity 28.5%; Pred. No. 0.014;
 Matches 47; Conservative 17; Mismatches 60; Indels 41; Gaps 5;

QY 105 EEEEEEFEVPAQNAKAVAKPAKPAVPAVDDEEDSDG-----MDE 150
 DB 221 KSSDERDKKAAAKASKNVAPTKKAASSDSEDEDEDAKPVSKPAVAKSKKPS 280
 QY 151 DSDGSEDESEETPKPKPAKRAETTPKAVSAKKA-----AVTPQ 196
 DB 281 SDDDDDDSSSDDKKPAKSKVEGRSDSDDDHKNNIDKSSDSDSESESDPEI 340

QY 197 KTDK-----KGGKAANSPKASQVSGCGCKTNTSGN 231
 DB 341 KTPQKKIKDVENVDAGKSGKAPNTPATPNETS-GS--KTFVGN 382

RESULT 15
 DMS

nucleolin - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
 C:Accession: A29958; A40769; A56240; I84688
 R:Bouton, H.M.; Lapeyre, B.; Amaritic, F.
 J. Mol. Biol. 200, 627-638, 1988
 A:Title: Structure of the mouse nucleolin gene. The complete sequence reveals that ea
 A:Reference number: A29958; MUID:88316930
 A:Accession: A29958
 A:Molecule type: DNA
 A:Residues: 1-707 <B0>
 A:Cross-references: GB:X07699; NID:g53453; PIDN:CAA30538.1; PID:g53454
 R:Pasternack, M.S.; Bieker, K.J.; McInerney, T.N.
 J. Biol. Chem. 266, 14703-14708, 1991
 A:Title: Granzyme A binding to target cell proteins. Granzyme A binds to and cleaves
 A:Reference number: A40769; MUID:91317840
 A:Accession: A40769
 A:Molecule type: Protein
 A:Residues: 2-20, 'X', 22-24 <PAS>
 R:Yang, T.H.; Tsai, W.H.; Lee, Y.M.; Lei, H.Y.; Lai, M.Y.; Chen, D.S.; Yeh, N.H.; Lee
 Mol. Cell. Biol. 14, 6068-6074, 1994
 A:Title: Purification and characterization of nucleolin and its identification as a t
 A:Reference number: A56240; MUID:94344117
 A:Accession: A56240
 A:Molecule type: Protein
 A:Residues: 2-19; 558-567 <YAN>
 R:Bouton, H.
 Gene 68, 73-84, 1988
 A:Title: Sequence and structure of the nucleolin promoter in rodents: Characterizatio
 A:Reference number: I84688; MUID:89121496
 A:Accession: I84688
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-44 <RES>
 A:Cross-references: GB:M22089; NID:g200111; PIDN:AAA39841.1; PID:g554246
 C:Comment: This housekeeping protein is involved in the synthesis, packaging, and mat
 C:Genetics:
 A:Introns: 6/3; 45/3; 211/1; 273/1; 302/1; 349/2; 391/1; 432/3; 484/1; 525/2; 566/1;
 C:Superfamily: nucleolin; ribonucleoprotein repeat homology
 C:Keywords: DNA binding; duplication; nucleus: phosphoprotein; RNA binding; transcrip
 F:310-375/Domain: ribonucleoprotein repeat homology <RNM1>
 F:311-316/Region: RNA-binding RNP2 motif
 F:349-356/Region: RNA-binding RNP1 motif
 F:396-458/Domain: ribonucleoprotein repeat homology <RNM2>
 F:397-402/Region: RNA-binding RNP2 motif
 F:431-438/Region: RNA-binding RNP1 motif
 F:488-551/Domain: ribonucleoprotein repeat homology <RNM3>
 F:489-494/Region: RNA-binding RNP2 motif
 F:524-531/Region: RNA-binding RNP1 motif
 F:570-634/Domain: ribonucleoprotein repeat homology <RNM4>
 F:571-576/Region: RNA-binding RNP2 motif
 F:607-614/Region: RNA-binding RNP1 motif

Query Match 12.0%; Score 153.5; DB 1; Length 707;
 Best Local Similarity 24.1%; Pred. No. 0.015;
 Matches 61; Conservative 31; Mismatches 94; Indels 67; Gaps 6;

QY 6 IEVSKGPTVTPGECIL-----HVSQSLGEC 35
 DB 48 IOKGKKKATTPKAAVYVQTKAAVTPKAAVTPGKAAVTPAKNITTPAVVPTP 107
 QY 36 NKGK-----EFVPLHVKNQNLVLTSTENIPOLFCDLVDEKFEFSLHWKGSYFV 90
 DB 108 GKGAQAQAKALVPTPGKGAATPAKGAKNKAKKEDSDEDEDEDDEDEDEED 167

Fri May 3 08:56:16 2002

us-09-645-337-6.rpr

Page 7

[illegible]

Search completed: May 2, 2002, 22:07:30
Job time: 6678 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:07:34 : Search time 36.81 Seconds
(without alignments)
257.710 Million cell updates/sec

Title: US-09-645-337-6

Perfect score: 1279

Sequence: 1 MEFWGIEVKSQKPVYTPPE.....TNSGNALSHKAKHAANK 245

Scoring table: BIOSUM62
Gapop 10.0 , Capext 0.5

Searched: 105224 seqs, 38719550 residues 105224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224	17.5	412	1	FKB4_SPOFR
2	191.5	15.0	357	1	FKB4_DROME
3	170.5	13.3	299	1	NFM_XENLA
4	170	13.3	694	1	NOCL_CHICK
5	169	13.2	294	1	NFM_CHICK
6	166.5	13.0	713	1	NOCL_MESAU
7	155	12.1	294	1	NPM_HUMAN
8	153.5	12.0	411	1	MP62_LYTRI
9	153.5	12.0	706	1	NOCL_MOUSE
10	153	12.0	706	1	NOCL_HUMAN
11	151.5	11.8	712	1	NOCL_RAT
12	150.5	11.7	281	1	HS32_DICDI
13	149.5	11.7	644	1	NFM_RABIT
14	149	11.6	292	1	NPM_RAT
15	145.5	11.4	704	1	NP14_RAT
16	144	11.3	292	1	NPM_MOUSE
17	139.5	10.9	411	1	FKB3_YEAST
18	138	10.8	699	1	NP14_HUMAN
19	137	10.7	590	1	YMW7_YEAST
20	136	10.6	650	1	NOCL_XENLA
21	131.5	10.3	731	1	NR11_COPCI
22	130	10.2	1982	1	CHDM_DROME
23	129.5	10.1	1085	1	IFH1_YEAST
24	128.5	10.0	1411	1	TCOF_HUMAN
25	128	10.0	639	1	SSTR_CATRO
26	128	10.0	891	1	MAZ3_SCHCO
27	127.5	10.0	361	1	FKB4_SCHCO
28	126.5	9.9	723	1	SSRP_DROME
29	126	9.9	217	1	KSI_HYDAT
30	126	9.9	226	1	BASP_BOVIN
31	126	9.9	450	1	CYL_PARDE
32	126	9.9	918	1	TF38_SCHCO
33	125.5	9.8	200	1	NOPL_XENLA

34	125	9.8	618	1	ORC2_DROME	Q24168 drosophila
35	124.5	9.7	810	1	NFM_BOVIN	077188 bos taurus
36	123.5	9.7	845	1	NFM_RAT	P12839 rattus norv
37	123.5	9.7	1359	1	ATRX_CAEEL	090760 caenorhabd
38	123	9.6	346	1	RS6_AEDAE	090761 aedes aegy
39	123	9.6	857	1	NFM_CHICK	P16053 gallus gall
40	122.5	9.6	677	1	UBF1_XENLA	P25979 xenopus lae
41	122.5	9.6	915	1	NFM_HUMAN	P07197 homo sapien
42	122	9.5	517	1	T2FA_HUMAN	P35269 homo sapien
43	122	9.5	1095	1	NEB1_RAT	035867 rattus norv
44	121.5	9.5	1020	1	NFM_HUMAN	P12036 homo sapien
45	121.5	9.5	1087	1	NFM_MOUSE	P19246 mus musculu

ALIGNMENTS

RESULT 1	FKB4_SPOFR	STANDARD:	PRT:	412 AA.
AC	Q26486:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	46 kDa FK506-binding nuclear protein (peptidyl-prolyl cis-trans isomerase) (PPIase) (EC 5.2.1.8).			
OS	Spodoptera frugiperda (Fall armyworm).			
OC	Insecta: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta: Pterygota: Neoptera: Endopterygota: Lepidoptera: Glossata: Ditrysia: Noctuidae, Noctuidae: Amphipyridae: Spodoptera.			
OX	NCBI_Taxid=7108;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95074110: PubMed=7527037:			
RA	Ainemi E.S., Fernandes-Ainemi T., Pomeranke K., Robertson N.M., Dudley K., Dubois G.C., Litwack G.;			
RT	*FKBP46, a novel 519 insect cell nuclear immunophilin that forms a protein-kinase complex.;			
RL	J. Biol. Chem. 269:30828-30834(1994).			
CC	-1- FUNCTION: BINDS TO, AND IS INHIBITED BY FK506 AND RAPAMYCIN. BINDS DOUBLE-STRANDED DNA IN VITRO. PPIASES ACCELERATE PROTEIN FOLDING. PEPTIDE BONDS IN OLIGOPEPTIDES.			
CC	-1- SUBCELLULAR LOCATION: Nuclear.			
CC	-1- PTM: PHOSPHORYLATED BY A NUCLEAR KINASE IN THE PRESENCE OF MG2+ AND ATP.			
CC	-1- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: U15038; AAA58962.1; -			
DR	HSSP: P20071; 1PKJ			
DR	InterPro: IPR001179; FKBP_PPIase.			
DR	PIfam: PF00254; FKBP_1.			
DR	PROSITE: PS00453; FKBP_PPIASE_1; FALSE_NEG.			
DR	PROSITE: PS00454; FKBP_PPIASE_2; 1.			
DR	PROSITE: PS0059; FKBP_PPIASE_3; 1.			
KW	Isomerase; Rotamase; Nuclear protein; DNA-binding; Phosphorylation.			
FT	DOMAIN 91..112			
FT	DOMAIN 120..145			
FT	DOMAIN 152..216			
FT	DOMAIN 219..302			
FT	DOMAIN 324..412			
SQ	SEQUENCE 412 AA; 45810 MW; F2A69159AER4FE22 CRC64;			

Query Match 17.5%; Score 224; DB 1; Length 412;

[illegible]

RESULT 2
PREDICTION

ID	PKR4_DROME	STANDARD;	PRT;	357	AA.
1	PKR4_DROME	STANDARD;	PRT;	357	AA.

DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 39 kDa FKS06-binding nuclear protein (Peptidyl-prolyl cis-trans
 isomerase) (PPIase) (EC 5.2.1.8).
 GN FKS06-BP1 OR FKBP39 OR CG6226.
 OS Drosophila melanogaster (Fruit fly).
 OC Euarthropala: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha
 OC Ephyridioidea: Drosophilidae: Drosophila.
 NCBI_TaxID:7227;

RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON-S.
 RX MEDLINE-95278752; PubMed-7538962;
 RA Theopold U., Dal Zotto L., Hüllmark D.;
 RT "KRP39, a Drosophila member of a family of proteins that bind the
 RT immunosuppressive drug FK506.";
 RL Immune 156:247-251(1995).

RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson S.C., Miklos G.L.G.,
RA Abil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Hallee R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
RA Borkova D., Borichan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke K., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Iodiscon K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottell J.B., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwu K.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.
RA Mekulov G., Bilshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relnett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkasz R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massarman D.A., Weinstock G.M., Weissendach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: PIPASIS ACCELERATE THE FOLDING OF PROTEINS, FKBP506
CC MAY FUNCTION IN A SIGNAL TRANSDUCTION CASCADE DURING EARLY
CC DEVELOPMENT.
CC
CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC
CC -1- TISSUE SPECIFICITY: UB1OUTOOLSY EXPRESSED, HIGHEST LEVELS IN
CC OVARY.
CC
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING ALL STAGES OF DEVELOPMENT
CC WITH HIGHEST EXPRESSION IN EARLY EMBRYO.
CC
CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PIASB FAMILY.
CC
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.gib.ch).
CC -----
CC EMBL: Z46894; CAAB6996.1; -.
DR EMBL: AF003708; AAP55172.1; AIT_SEQ.
DR HSSP: P27124; IROT.
DR FlyBase: FBgn0013269; FK506-bp1.
DR InterPro: IPR001179; FKBP_PPIase.
DR Pfam: PF00254; FKBP. 1.
DR PROSITE: PS00453; FKBP_PPIASE. 2; 1.
DR PROSITE: PS00454; FKBP_PPIASE. 2; 1.
DR PROSITE: PS00559; FKBP_PPIASE.3; 1.
DR PROSITE: PS00559; FKBP_PPIASE.3; 1.
KW Isomerase; Notamase; Nuclear protein.
FT DOMAIN 89
FT DOMAIN 119 183 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 186 247 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 269 357 LYS-RICH (BASIC).
FT DOMAIN 357 357 PPIASE, FKBP-TYPE.
FT CONFILCT 187 187 A -> R (IN REF. 1).
SQ SEQUENCE 357 AA; 39343 MW; EFOA87B31738BB30 CNG64;

Query Match 15.08; Score 191.5; DB 1; Length 357;

Matches	72;	Conservative	42;	Mismatches	110;	Indels	59;	Gaps	10;
---------	-----	--------------	-----	------------	------	--------	-----	------	-----

```

QY 3 FVGLEWKSGRVYVTPREGILLIHSQASLGECKNNKGEFPLHVKVGNQNLVLGTSTEN 62
Db 4 FVGNNMKPERKYSOTILKS--FHISGVAL-----DKGGEALYLAARKEQYIATV--TKA 55
QY 63 IPOLFCDLVPLPK--PEPLSHTWKGSVYFVGVKTTNIEPQCSSEEEEEEEFVPAGNNAKA 122
Db 56 IPOVALDLNFKSGRIMFYTAGDASVSLGVL-----HDISEDDDDDDQMTIENILNS 109
QY 122 VAKPRAKAEVKPAVD-----EEDSDSD-----GMEDESDSGEDSE 155
Db 110 KAIRNKSKE-----DDEDESGESEDDEDDDDSOQIIIEYSEFLENGEEFDDDDVDEDN 164
QY 160 EEPFPKPKPASKKRANNTTKPARY-----SAKKA-----KVAVTPQKTBKK 202

```


Db 165 ESSGEDEDDSDSEEEOPAKAVAKLSPGASAKSGKEONGVAKKEAKOAKKEKP 224
 QY 203 KCGKAAOSPKSASQVSGSCCKTTPNSGMALESHNKAHAAAK 245
 Db 225 EAKKEQPAKKEPAKQOPASKDPRTITGCGYKIVDYVVGKEEAK 267

RESULT 3

NPM_XENLA STANDARD: PRT: 299 AA.
 ID NPM_XENLA
 AC P07222:
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin)
 DE (Nucleolar protein NO38).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=83355;
 RN 11
 RP SEQUENCE FROM N.A.

RC TISSUE-Ovary;
 RX MEDLINE=86004394; PubMed=3308448;
 RA Schmidt-Zachmann M.S., Huegle-Doerr B., Franke W.W.;
 RT "A constitutive nucleolar protein identified as a member of the
 RT nucleoplasm family."
 RL EMBL J. 6:1881-1890(1987).

CC -1- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
 AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
 CC ASSEMBLY AND/OR TRANSPORT OF RIBOSOME (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEOLASMIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).

DR EMBL: X05496; CAA29046.1; -
 DR PIR: A29681; A29681.
 DR InterPro: IPR004301; Nucleoplasmn.

DR Pfam: PF03066; Nucleoplasmn; 1.
 KW Nuclear protein; Phosphorylation; RNA-binding; Repeat.

FT DOMAIN 1 7
 FT DOMAIN 122 137 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 153 158 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 160 187 ASP/GLU-RICH (HIGHLY ACIDIC).
 FT DOMAIN 189 195 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 218 242 4 x 3 AA REPEATS OF K-T-P.
 FT REPEAT 218 220
 FT REPEAT 221 223
 FT REPEAT 237 239
 FT REPEAT 240 242
 FT REPEAT 240 242
 SQ SEQUENCE 299 AA: 33532 MW: C16CDP9565090843 CRC64:

Query Match 13.3%; Score 170.5; DB 1; Length 299;
 Best Local Similarity 31.5%; Pred. No. 0.00025;
 Matches 56; Conservative 20; Mismatches 67; Indels 35; Gaps 7;

QY 83 GKGSTVFGYKTPNIEPGYSEEEEEEVAGNAKAAVAPKPAVKPAVDEDEDE 142
 Db 107 GSGPYVVSQHLVLELDSDEDEDEHPSPK-NAKRIAPDSASVPRKTRLEEEED 165
 QY 143 SD-SQGMDDDDSDGSEFEFEFP-KKPASSKKRA-----NETTPRAPS 185
 Db 166 SDEDDDDDDDEDEDEEETPVKKTSTSKAAQKLNHNKASALSTQTKPTPPO 225

QY 186 AKK-----AKVAVTPQKTDK-----KCGKAAOSPKSASQVSGSCCKTTPN 228
 Db 226 KKGQDTKTPQKTPKTPKPLISSEELAKKQMTYLEKGVLPKVEKANYVK--NCRFTEN 281

RESULT 4

NUCL_CHICK STANDARD: PRT: 694 AA.
 ID NUCL_CHICK
 AC P15771;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nucleolin (Protein C23).
 DE Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN 11
 RP SEQUENCE FROM N.A.

RX MEDLINE=90206792; PubMed=2320420;
 RA Maridori G., Nigg E.A.;
 RT "cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major
 RT nucleolar proteins."
 RL Nucleic Acids Res. 18:1286-1286(1990).

CC DISCUSSION OF SEQUENCE.
 RX MEDLINE=90304215; PubMed=2114180;
 RA Maridori G., Nigg E.A.;
 RT "Structure and developmental expression of chicken nucleolin and
 RT NO38: coordinate expression of two abundant non-ribosomal nucleolar
 RT proteins."
 RL Biochim. Biophys. Acta 1049:126-133(1990).

RP SEQUENCE OF 407-694 FROM N.A.
 RX MEDLINE=89119560; PubMed=2914325;
 RA Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;
 RT "Major nucleolar proteins shuttle between nucleus and cytoplasm."
 RL Cell 56:379-390(1989).

CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
 CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
 CC CHROMATIN AND PERIRIBOSOMAL PARTICLES. IT INDICES CHROMATIN
 CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
 CC ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
 CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).

DR EMBL: X17199; CAA35060.1; -
 DR EMBL: M21791; AAA48983.1; -
 DR PIR: S08414; DNCHNL.
 DR HSSP: P09651; 10P1.
 DR InterPro: IPR000504; RRM.
 DR Pfam: PF00076; RRM; 4.
 DR SMART: SM00360; RRM; 4.
 DR PROSITE: PS00102; RRM; 4.
 DR PROSITE: PS00030; RRM_RNP_1; 3.
 KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
 KM Nuclear protein; Phosphorylation.
 FT DOMAIN 124 141 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 170 192 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 217 247 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 281 357 RNA-BINDING (RRM) 1.
 FT DOMAIN 371 445 RNA-BINDING (RRM) 2.
 FT DOMAIN 461 535 RNA-BINDING (RRM) 3.

RT hyperphosphorylated protein in a Ki-1 lymphoma cell line with
 RT chromosomal translocation t(2;5)(p21;p11).
 [9]
 RN SEQUENCE OF 1-133 FROM N.A.
 RP TISSUE-BONE MARROW;
 KC MEDLINE:96151966; PubMed-8562957;
 KA Redner R.L., Rush E.A., Paas S., Rudert W.A., Corey S.J.;
 RT "The t(5;17) variant of acute promyelocytic leukemia expresses a
 RT nucleophosmin-retinoic acid receptor fusion.";
 RL Blood 87:882-886(1996).
 RN [10]
 RP SEQUENCE OF 115-134.
 RX MEDLINE:86111864; PubMed-3944116;
 RA Chan P.-K., Aldrich M.B., Cook R.G., Busch H.;
 RT "Amino acid sequence of protein B23 phosphorylation site.";
 RL J. Biol. Chem. 261:1868-1872(1986).
 RN [11]
 RP SEQUENCE OF 33-54.
 KC TISSUE-Colon carcinoma;
 RX MEDLINE:97295306; PubMed-9150948;
 RA J.I.H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;
 RT "A two-dimensional gel database of human colon carcinoma proteins.";
 RL Electrophoresis 18:605-613(1997).
 RN [12]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE:96152893; PubMed-8570204;
 RA Yoneda-Kato N., Look A.T., Kirtstein M.N., Valentine M.B.,
 RA Raimondi S.C., Cohen K.J., Carroll A.J., Morris S.W.;
 RT "The t(3;5)(q25;q34) of myelodysplastic syndrome and acute myeloid
 RT leukemia produces a novel fusion gene, NPM-MPL.";
 RL Oncogene 12:265-275(1996).
 CC -1- FUNCTION: ASSOCIATED WITH NUCLEOLAR RIBONUCLEOPROTEIN STRUCTURES
 CC AND BIND SINGLE-STRANDED NUCLEIC ACIDS. IT MAY FUNCTION IN THE
 CC ASSEMBLY AND/OR TRANSPORT OF RIBOSOME.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMERS UNDER CERTAIN CONDITIONS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. GENERALLY NUCLEOLAR, BUT IS
 CC TRANSLATED TO THE NUCLEOLAR PLASMA IN CASE OF SERUM STARVATION OR
 CC TREATMENT WITH ANTICANCER DRUGS.
 CC -1- PTM: PHOSPHORYLATED.
 CC -1- DISEASE: A FORM OF NON-HODGKIN'S LYMPHOMA IS CHARACTERIZED BY A
 CC CHROMOSOMAL TRANSLOCATION T(2;5)(p23;q35) THAT INVOLVES NPM AND
 CC ALK.
 CC -1- DISEASE: A FORM OF ACUTE PROMYELOCYTIC LEUKEMIA IS CHARACTERIZED
 CC BY A CHROMOSOMAL TRANSLOCATION T(5;17)(q32;q11) THAT INVOLVES NPM
 CC AND RARA.
 CC -1- DISEASE: PARTICIPATES IN A T(3;5)(q25;q34) CHROMOSOMAL
 CC TRANSLOCATION. THE PRODUCT, A NPM-MPL FUSION PROTEIN, IS
 CC RESPONSIBLE FOR MYELODYSPLASTIC SYNDROME (MDS) WHICH PROGRESSES TO
 CC ACUTE MYELOID LEUKEMIA (AML).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEOPHOSMIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M23613; AAA36380.1; -
 DR EMBL: M28699; AAA58386.1; -
 DR EMBL: M26697; AAA36385.1; -
 DR EMBL: 089321; AAB94739.1; -
 DR EMBL: 089309; AAB94739.1; JOINED.
 DR EMBL: 089310; AAB94739.1; JOINED.
 DR EMBL: 089311; AAB94739.1; JOINED.
 DR EMBL: 089313; AAB94739.1; JOINED.
 DR EMBL: 089314; AAB94739.1; JOINED.
 DR EMBL: 089317; AAB94739.1; JOINED.
 DR EMBL: 089319; AAB94739.1; JOINED.
 DR EMBL: X16934; CAA34809.1; -
 DR EMBL: J02590; AAA36473.1; -
 DR EMBL: J02590; AAA36473.1; -

DR EMBL: M31004; AAA36474.1; ALT_SEQ.
 DR EMBL: 004946; AAA58698.1; ALT_TERM.
 DR EMBL: D45915; BAA08343.1; ALT_TERM.
 DR EMBL: 041742; AAB00112.1; ALT_TERM.
 DR EMBL: 041743; AAB00113.1; ALT_TERM.
 DR PIR: A30137; A30137.
 DR PIR: A32915; A32915.
 DR PIR: A26080; A26080.
 DR PIR: B26080; B26080.
 DR PIR: S06926; S06926.
 DR PIR: A33423; A33423.
 DR SWISS-2DPAGE; P06748; HUMAN.
 DR MIM: 164040; -
 DR InterPro: IPR004301; Nucleoplasm.
 DR Pfam: PF03066; Nucleoplasm.
 KW Nuclear protein; Phosphorylation; RNA-binding; Proto-oncogene;
 RN Chromosomal translocation.
 FT DOMAIN 1 9 MET-RICH.
 FT DOMAIN 120 132 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 152 157 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 161 188 ASP/GLU-RICH (HIGHLY ACIDIC).
 FT DOMAIN 191 197 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT MOD_RES 125 125 PHOSPHORYLATION.
 FT MOD_RES 199 199 PHOSPHORYLATION.
 FT SITE 175 176 BREAKPOINT FOR TRANSLATION TO FORM NPM-MPL.
 FT CONFLICT 129 129 E -> D (IN REF. 10).
 FT CONFLICT 213 213 D -> P (IN REF. 6).
 FT CONFLICT 216 216 P -> S (IN REF. 6).
 FT CONFLICT 219 221 TPR -> SSS (IN REF. 6).
 SQ SEQUENCE 294 AA; 32575 MW; 620BC7BA2EA40054 CMC64;
 Query Match 12.1%; Score 155; DB 1; Length 294;
 Best Local Similarity 28.3%; Pred. No. 0.0023;
 Matches 43; Conservative 25; Mismatches 56; Indels 28; Gaps 5;
 QY 83 GKGSYFVGYCTKPNIPEOCYSEEEER-----EYVQNAKAAVAKPKAPVAK 133
 DB 105 GSGPYHISQHLVAVEAESEDEEEEDYKLSIGKRSAPGG-----SKVPQKKVLA 159
 QY 134 PAVDDDESDSDGDDEDDSDGDESEDEEPPPK-----PASSKRRANETTPKAPVAK 187
 DB 160 ADEDDDDDEDEDEDDDDDDDFDDEAEKAPVKSIRDTPAKNAKSNQ-----NGK 212
 QY 188 KAKVAVTPQ-KTDEKKKGKGAANSPKSAQV 218
 DB 213 DSKPSTPRKSGSPKKGKPKTPKPCSSV 244
 RESULT 8
 ID MP62_LYTP1 STANDARD; PRT; 411 AA.
 AC P91753;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Mitotic apparatus protein p62.
 OS Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Echinoidae; Echinacea; Temnopleuroidea; Tokopneustidae;
 OC Lytechinus.
 OX NCBI_TaxId=7653;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:97166213; PubMed-9013612;
 RA Ye X., Sloboda R.D.;
 RT "Molecular characterization of p62, a mitotic apparatus protein
 RT required for mitotic progression.";
 RL J. Biol. Chem. 272:3606-3614(1997).
 CC -1- FUNCTION: REQUIRED FOR MITOTIC PROGRESSION. BINDS TO CHROMATIN.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: PHOSPHORYLATED BY CAM-KINASE II IN VITRO.


```

Db 107 GKKGAQAQAKALVPTPKKGAATPAKCAKKNKAKKEDSDPRHEDHDSDEDEDEED 166
QY 91 GYKTPMT-----EPGYSEEEEEEEVPAKNAKAV-----AKPAK 128
Db 167 EEPEPIYGVKPAKAAAPAPASDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 226
QY 129 PAEYKPAV-----DDEDESDSDGDM-EDDSOGHSEEEPPPKKASKKANET 178
Db 227 PAVVPAKNAKSAVAEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 286
QY 179 TPAPVSAKRAK 191
Db 287 KKKAPPAKAKK 299

RESULT 10
NUCL_HUMAN STANDARD: PRT: 706 AA.
AC P19338;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
NCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89290043; PubMed=2737305;
RA Stivastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT Cloning and sequencing of the human nucleolin cDNA.;
RL FEBS Lett. 250:99-105(1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368666; PubMed=2394707;
RA Stivastava M., McArtide O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT "Genomic organization and chromosomal localization of the human
RT nucleolin gene.";
RL J. Biol. Chem. 265:14922-14931(1990).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M60856; AA59954.1; -
DR PIR: A35804; A35804.
DR PIR: S04631; S04631.
DR HSSP: P09651; 1HA1.
DR Aarhus/Chemt-2DPACK; 1210; NHPHCF.
DR MIM: 164035; -
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; RRM; 4.
DR SMART: SM00360; RRM; 4.
DR PROSITE: PS0102; RRM; 4.
DR PROSITE: PS00030; RRM_RNP_1; 3.
KM Nuclear protein. Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding.
FT INIT_MET 0 0
FT DOMAIN 142 170 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 184 208 ASP/GLU-RICH (ACIDIC).

```

```

FT DOMAIN 233 270 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 306 382 RNA-BINDING (RRM) 1.
FT DOMAIN 392 465 RNA-BINDING (RRM) 2.
FT DOMAIN 485 559 RNA-BINDING (RRM) 3.
FT DOMAIN 571 643 RNA-BINDING (RRM) 4.
FT DOMAIN 645 694 ARG/CLY/PHE-RICH.
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT REPEAT 57 64 X-X.
FT REPEAT 74 81 1.
FT REPEAT 82 89 2.
FT REPEAT 90 97 3.
FT REPEAT 98 103 4.
FT REPEAT 104 111 5 (INCOMPLETE).
FT REPEAT 119 126 6.
FT REPEAT 127 134 7.
FT MOD_RES 144 144 8.
FT MOD_RES 152 152 8.
FT MOD_RES 183 183 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 183 183 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 706 AA; 76213 MW; 85A2F2CA22FA03DB CRC64;

Query Match 12.0%; Score 153; DB 1; Length 706;
Best Local Similarity 37.1%; Pred. No. 0.0077;
Matches 39; Conservative 9; Mismatches 47; Indels 10; Gaps 2;

QY 104 EEEEEEEVPAKNAKAVAKPAKPAVAVDE-----EDSDSDGMDSDGSD 157
Db 200 DEEDSEEEAMETTPAKGKAKKAVPVKAKVAEDEDDEDEDEDEDEDEDEDEDEDE 259
QY 158 -----SEEEPTPKKPAKSKKRAKPAKPAKPAKPAKPAKPAKPAKPAKPAK 198
Db 260 DEEEEEEEEPKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAKPAK 304

RESULT 11
NUCL_RAT STANDARD: PRT: 712 AA.
AC P13383;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
NCL OR NUC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269607; PubMed=2347493;
RA Bourbon H.-M., Amalric F.;
RT "Sequence and structure of the nucleolin promoter in rodents:
RT characterization of a strikingly conserved CpG island.";
RL Gene 68:73-84(1988).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M55022; AAA41732.1; -
DR EMBL: M55015; AAA41732.1; JOINED.
DR EMBL: M55017; AAA41732.1; JOINED.
DR EMBL: M55020; AAA41732.1; JOINED.
DR EMBL: M22090; AAA41733.1; -
DR PIR: JH0148; JH0148.
DR HSP: P09651; 1HA1.
DR InterPro: IPR00504; RRM.
DR Pfam: PF00076; rrm; 4.
DR SMART: SM00360; RRM; 4.
DR PROSITE: PS00102; RRM; 4.
DR PROSITE: PS00030; RRM_RNP_1; 3.
KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding.
FT INIT_MET 0 0
FT DOMAIN 142 167 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 187 215 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 241 274 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 310 366 RNA-BINDING (RRM) 1.
FT DOMAIN 396 469 RNA-BINDING (RRM) 2.
FT DOMAIN 488 562 RNA-BINDING (RRM) 3.
FT DOMAIN 574 649 RNA-BINDING (RRM) 4.
FT DOMAIN 651 702 ARG/GLY/PHE-RICH.
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-X-X.
FT REPEAT 57 64
FT REPEAT 74 81
FT REPEAT 82 89
FT REPEAT 90 97
FT REPEAT 98 103
FT REPEAT 104 111
FT REPEAT 119 126
FT REPEAT 127 134
SQ SEQUENCE 712 AA; 77016 MW; 68774A21AE50F90 CRC64;

Query Match 11.8%; Score 151.5; DB 1; Length 712;
Best Local Similarity 24.6%; Pred. No. 0.0096;
Matches 65; Conservative 24; Mismatches 88; Indels 87; Gaps 8;

QY 6 IEKSGKPTVTPPEEGLH-----VQASLGECK 35
DB 47 IPQKGRKATTTTPAKKVVVSOTKKAAPTPPAKKAATVPKKAATVPKAVVPP 106
QY 36 NKKG-----EFVPIHVKVGNQNLVLTSTENIPQLFCDLVFDK-----FF 76
DB 107 GKGAQAQALAVLTPGKKGAVTPPAKAKNGKAKKEDSDEDEDEDEDEDEDEDE 166
QY 77 ELSHTGKGSVTVGKRTNIEPGYSEEE-----EEVPAGNAKAV-- 122
DB 167 E-----PPVVKGVKPAKAPAPASDEDEDEDEDEDEDEDEDEDEDEEEVE 218
QY 123 -----AKPKAPAEVPA-----VDDEDESDSGMDDESDSGSESEEPKPK 167
DB 219 ITPAKGGKTPAKAVPVKAKASVAEEDEDEDEDEDEDEDEDEDEDEDEEEEP 277
QY 168 PASSKRRANETTPKAPVSAKAKAV 191
DB 278 AARGKRRKEMTKQKAPAKKOKI 301
RESULT 12
HS32.DICDI
ID HS32.DICDI STANDARD: PRT; 281 AA.
AC P54658;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE 32 kDa heat shock protein (4-1 protein).
GN HSPC OR HSP32.
OS Dictyostelium discoideum (Slime mold).
CC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
CC NCBI_TaxID=44689;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96001265; PubMed=7557471;
RA de Maria A.C., Gomes S.L., Juliana M.H., Mazzarella R., Klein C.;
RT *Cloning of a cDNA encoding a novel heat-shock protein from
RT Dictyostelium discoideum.*
RL Gene 163:163-164(1995).
CC -I- DEVELOPMENTAL STAGE: PRESENT AT HIGH LEVELS IN GROWING CELLS BUT
CC DECREASES DRAMATICALLY DURING THE EARLY HOURS OF DEVELOPMENT.
CC -I- INDUCTION: BY HEAT SHOCK.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U40211; AAC47710.1; -
DR EMBL: L39778; AAA9510.1; -
DR Dictydb: DD01076; hspc.
KW Heat shock.
FT DOMAIN 251 257 POLY-GLN.
FT DOMAIN 262 265 POLY-ASN.
SQ SEQUENCE 281 AA; 31481 MW; 2C64F57C847D7CA4 CRC64;

Query Match 11.8%; Score 150.5; DB 1; Length 281;
Best Local Similarity 23.4%; Pred. No. 0.0042;
Matches 66; Conservative 47; Mismatches 94; Indels 75; Gaps 11;

QY 1 MEFGIEVKSGRVTVTPPEEGLIHVSQASLGECKNKG-EFVPIHVKVGNQNLVLTSTENIPQLFCDLVFDK-----FF 76
DB 2 MGFQGIIVKEPVNLEDESDIFHLTKAIHPRKQSGKYTLNAVLSLMEDEKEEDV 61
QY 53 -----LVLTSTENIPQLFCDLVFDK-----KEELSHTWKGSVTVGKRTN 95
DB 62 DDEESPREDIIVEIPGILEACKIQIDILNHYNGQIVRELAQEMNAGYVALSGSVI 121
QY 96 NIEPGYSEEEEEEVEVPAGNAKAVAKPKKAPAEVPAVDEDEDESDSGMDDESDG 155
DB 122 TMEQGCDEDCDDHCT-----NHEDEDEIDSDER-FGSDODE 160
QY 156 EDSSEEE-----PTPK-----PASSKRRANETTPKAPVSAKAKAVATPQKTD 200
DB 161 EDSDEDEIPQLAPATPKKGTIEISEVPSKKEKTPPE--PKVPPPKKEGV--KQTPQ 215
QY 201 KKKGKAANQSPKASQVSGSGCKTFNSGNALSHNKAHA 242
DB 216 QOK--KAAAOQPEKAN-----NKPAAASPAKPONOSKNA 248
RESULT 13
NFM_RABIT
ID NFM_RABIT STANDARD: PRT; 644 AA.
AC P54938;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet M protein (160 kDa neurofilament protein)
DE (Neurofilament medium polypeptide) (NF-M) (Fragment).
GN NEF3 OR NEFM OR NFM.
OS Oryctolagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CC NCBI_TaxID=9986;
RN [1]

FT AH (IN ISOFORM R23.2).
 SQ SEQUENCE 292 AA; 32560 MW; 1372A474F9ED2457 CRC64;
 Query Match 11.68; Score 149; DB 1; Length 292;
 Best Local Similarity 27.98; Pred. No. 0.0055;
 Matches 39; Conservative 26; Mismatches 69; Indels 6; Gaps 2;
 OY 83 GKGSYFYGYTPTNIEPGYSEEESEEEVPAAGNAKAKPKAPVAVPAVDEEDE 142
 DB 105 GSGPHVHSGUHLVAEEAESEDEDEYKLLGSGKASAGGKVKYQKKVLDDEDE 164
 OY 143 SDSGMDSDSDSDSESESEEPKPKASKKRANET---TPKAPVSAKAKAVAVTPQKT 198
 DB 165 DDEDEDEDEDEDEDEDEETEEKVPV--KKSVRTPAKKNQKSNQNGKDLKPTPSKSG 222
 OY 199 DEKKKGKAAANOSPKASASQV 218
 DB 223 QESFKKOKERTPKTPKGPSSV 242

RESULT 15

NP14.RAT . STANDARD; PRT; 704 AA.

AC P41777;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa

DE Nucleolar phosphoprotein) (Noppl40) (Nucleolar and coiled-body

DE phosphoprotein 1).

GN NOLC1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.

RC TISSUE=Liver;

RX MEDLINE=92323542; PubMed=1623516;

RA Meier U.T., Blobel G.;
 "Noppl40 shuttles on tracks between nucleolus and cytoplasm.";

RT Cell 70:127-138(1992).

RL [2]

RP INTERACTION WITH NOPS AND FIBRILLARIN.

RX MEDLINE=20143579; PubMed=10679015;

RA Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;

"Conserved composition of mammalian box H/ACA and box C/D small

nucleolar ribonucleoprotein particles and their interaction with the

common factor Noppl40.";

RT Mol. Biol. Cell 11:567-577(2000).

RL -1- FUNCTION: RELATED TO NUCLEOGENESIS. MAY PLAY A ROLE IN THE

MAINTENANCE OF THE FUNDAMENTAL STRUCTURE OF THE FIBRILLAR CENTER

AND DENSE FIBRILLAR COMPONENT IN THE NUCLEOLUS. IT HAS INTRINSIC

GTPASE AND ATPASE ACTIVITIES. MAY PLAY AN IMPORTANT ROLE IN

TRANSCRIPTION CATALYZED BY RNA POLYMERASE I (BY SIMILARITY).

-1- SUBUNIT: Interacts with Dkc1/Nap57, Nops/Nap65 and fibrillarin.

-1- SUBCELLULAR LOCATION: SHUTTLES ON CURVILINEAR TRACKS BETWEEN

NUCLEOLUS AND CYTOPLASM. THESE TRACKS EXTEND FROM THE DENSE

FIBRILLAR COMPONENT OF THE NUCLEOLUS ACROSS THE NUCLEOLAR PLASM TO

A LIMITED NUMBER OF NUCLEAR PORE COMPLEXES.

-1- PTM: THIS PROTEIN UNDERGOES RAPID AND MASSIVE PHOSPHORYLATION AND

DEPHOSPHORYLATION ON CR-II AND PKC SITES. NOPPL40 IS ONE OF THE

MOST PHOSPHORYLATED PROTEINS IN THE CELL.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

DR EMBL; M94287; AAA41718.1; -.
 DT EMBL; M94288; AAA41719.1; -.
 KW Nucleolar protein; Phosphorylation; Repeat; GTP-binding; ATP-binding.
 FT DOMAIN 84 570
 FT REPEAT 84 95
 FT REPEAT 127 138
 FT REPEAT 170 181
 FT REPEAT 231 242
 FT REPEAT 274 285
 FT REPEAT 335 346
 FT REPEAT 373 384
 FT REPEAT 434 445
 FT REPEAT 479 490
 FT REPEAT 524 535
 FT REPEAT 559 570
 FT MOD_RES 567 567
 FT VARIANT 150 150
 SQ SEQUENCE 704 AA; 73562 MW; 14DF1BF2DE483EA3 CRC64;

Query Match 11.48; Score 145.5; DB 1; Length 704;

Best Local Similarity 25.08; Pred. No. 0.023;

Matches 52; Conservative 22; Mismatches 65; Indels 69; Gaps 5;

OY 103 SEEESEEEVPAAGNAKAVAKPKAPVAVPAVDEDESDGMDSDGDESEEE 162
 DB 339 SDSSESEKKTP---AKTVSKTTPAKPAKYKKAESSSDSDSDDEPAKPVSAITKS 395

OY 163 P-----TPKPA-----SKKRAANE--TTPK 181

DB 396 PLKRPVAVTPPKPAKAVATPKOPAGSGOKPQSRKADSSSESESESEFEATKKSVTTPK 455

OY 182 APVSAKKA-----KVAVTPQKTDEKKKGKGAANSPKAS 216

DB 456 ARVTAKAAPSLPAKQAPRAGDSSSDSESESESEKTPPKPAKKAACAAPKPPVK 515

OY 217 QVSGCGCKTFNSGNALSHNKAHAA 244

DB 516 KAAAESSSSSSSDSESESEKPKKSKA 543

Search completed: May 2, 2002, 22:31:08
 Job time: 1414 sec

• • *

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:03:59 ; Search time 113.15 Seconds

(without alignments)
374.580 Million cell updates/sec

Title: US-09-645-337-6

Perfect score: 1279
Sequence: 1 MEFWGIEVSKGKPVTVPEE.....TENGNALESHNKAKHAAK 245

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1279	100.0	245	10	Q9FVE6	Q9FVE6 arabidopsis
2	1274	99.6	257	10	Q22238	Q22238 arabidopsis
3	600.5	47.0	306	10	Q9FN06	Q9FN06 arabidopsis
4	594	46.4	305	10	Q49209	Q49209 arabidopsis
5	530.5	41.5	303	10	Q9M4U5	Q9M4U5 zea mays (m
6	513.5	40.1	294	10	Q9M4R4	Q9M4R4 arabidopsis
7	511.5	40.0	297	10	Q9M4R5	Q9M4R5 oryza sativ
8	498.5	39.0	296	10	Q9LZR5	Q9LZR5 arabidopsis
9	493	38.5	300	10	Q9M4U4	Q9M4U4 zea mays (m
10	482	37.7	307	10	Q24591	Q24591 zea mays (m
11	324.5	25.4	286	10	Q94F81	Q94F81 zea mays (m
12	306	23.9	203	10	Q9M4R3	Q9M4R3 arabidopsis
13	179	14.0	477	10	Q93ZG9	Q93ZG9 arabidopsis
14	165.5	12.9	990	13	Q91803	Q91803 xenopus lae
15	164	12.8	487	10	Q9STK2	Q9STK2 arabidopsis
16	162.5	12.7	707	11	Q99K50	Q99K50 mus musculu

17	158.5	12.4	399	11	Q99L14	Q99L14 mus musculu
18	158.5	12.4	513	5	Q95P08	Q95P08 chironomus
19	157	12.3	294	4	Q96AT6	Q96AT6 homo sapien
20	157	12.3	346	5	Q9NLA3	Q9NLA3 asterina pe
21	157	12.3	1110	13	Q91255	Q91255 petromyzon
22	155.5	12.2	444	11	Q9CT46	Q9CT46 mus musculu
23	155	12.1	179	4	Q9NMX3	Q9NMX3 homo sapien
24	155	12.1	214	4	Q9BRT9	Q9BRT9 homo sapien
25	155	12.1	259	4	Q9BYG9	Q9BYG9 homo sapien
26	155	12.1	294	4	Q96EA5	Q96EA5 homo sapien
27	153.5	12.0	265	4	Q96DC4	Q96DC4 homo sapien
28	153.5	12.0	635	10	Q40363	Q40363 medicago sa
29	151.5	11.8	715	11	Q9Q2X1	Q9Q2X1 rattus norv
30	150	11.7	606	4	Q9Y2D9	Q9Y2D9 homo sapien
31	148	11.6	296	13	Q91802	Q91802 xenopus lae
32	146.5	11.5	798	5	Q45181	Q45181 caenorhabdi
33	145.5	11.4	312	5	Q26711	Q26711 trypanosoma
34	144	11.3	257	11	Q9DAY9	Q9DAY9 mus musculu
35	144	11.3	285	11	Q9DAV3	Q9DAV3 mus musculu
36	144	11.3	507	3	Q00880	Q00880 haematocyt
37	142.5	11.1	305	3	Q9C2P7	Q9C2P7 neurospora
38	140.5	11.0	649	10	Q9SU25	Q9SU25 arabidopsis
39	140.5	11.0	798	13	Q90307	Q90307 carassius a
40	140.5	11.0	1881	16	Q9L7Q2	Q9L7Q2 streptococ
41	138.5	10.8	688	5	Q9VNX6	Q9VNX6 drosophila
42	138.5	10.8	971	5	Q9XVS4	Q9XVS4 caenorhabdi
43	138	10.8	418	4	Q96J17	Q96J17 homo sapien
44	138	10.8	700	4	Q9BUV3	Q9BUV3 homo sapien
45	138	10.8	705	13	Q06459	Q06459 xenopus lae

ALIGNMENTS

RESULT 1
Q9FVE6 PRELIMINARY; PRT; 245 AA.
AC Q9FVE6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
DE HD2A.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_taxid=3702;
RM [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologs in
Arabidopsis thaliana.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF195545; AAC8472.1; -;
DR InterPro: IPR000822; Znf-C2H2.
DR SMART: SM00355; Znf-C2H2; 1.
DR PROSITE: PS00026; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 245 AA; 26372 MW; EBE4F7E0AD24F28 CRC64;

Query Match 100.0%; Score 1279; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 7.6e-89;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFWGIEVSKGKPVTVPEEGILIHVSQASIGECNNKKGEPVPLHVKKGNLVLTST 60
DB 1 MEFWGIEVSKGKPVTVPEEGILIHVSQASIGECNNKKGEPVPLHVKKGNLVLTST 60
QY 61 ENIPOLFCDLVFDEKFEFLSHTWGKGSYFVGKTPNIEPOGSEEEEEEVEPAGNAAK 120

```

|||||
Db 61 ENIPOLCDLVDFKPEFLSHTWKGSYFVGYKTPNIEPOGYSEEEEPKPKKRAKETTP 120
OY 121 AVAKPKAKPAEVPKPAVDEDESDSDCMDDDDSDGDESEFEPTPKKPASSKRRANETTP 180
Db 121 AVAKPKAKPAEVPKPAVDEDESDSDCMDDDDSDGDESEFEPTPKKPASSKRRANETTP 180
OY 181 KAPVSAKAKAVYVTPKOTDEKKKGAANOSPKSASQVSCGCKTTFNSGNALSHNKAK 240
Db 181 KAPVSAKAKAVYVTPKOTDEKKKGAANOSPKSASQVSCGCKTTFNSGNALSHNKAK 240
OY 241 HAAAK 245
Db 241 HAAAK 245

RESULT 2
022238 PRELIMINARY; PRT: 257 AA.
AC 022238;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN T32N15.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. COLUMBIA;
RA de la Bastide M.R., Parnell L.D., Kaplan N., Gnoj L., Hameed A.,
RA Schurz K., Hasegawa A., Gottesman T., Shohdy N., Granat S., Jensen K.,
RA Johnson A.F., Lodhi M., Dedina M., Martienssen R., McCombie W.R.,
RA "A. thaliana BAC T32N15 from chromosome V.";
RA Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: ACC002534; AAB70032.1;
DR InterPro: IPR000822; Znf-C2H2.
DR Pfam: PF00096; Znf-C2H2.1.
DR SMART: SM00355; Znf-C2H2.1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 257 AA; 27698 MW; C44043876689D5 CRC64;

```

Query Match 99.6%; Score 1274; DB 10; Length 257;
 Best Local Similarity 100.0%; Pred. No. 1.9e-88;
 Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MEFWGLGVKSGKPYVTPREBGLIHVSQASLGECNNKKGFPVPLHVKVGNOMLVLTIST 60
Db 1 MEFWGLGVKSGKPYVTPREBGLIHVSQASLGECNNKKGFPVPLHVKVGNOMLVLTIST 60
OY 61 ENIPOLCDLVDFKPEFLSHTWKGSYFVGYKTPNIEPOGYSEEEEPKPKKRAKETTP 120
Db 61 ENIPOLCDLVDFKPEFLSHTWKGSYFVGYKTPNIEPOGYSEEEEPKPKKRAKETTP 120
OY 121 AVAKPKAKPAEVPKPAVDEDESDSDCMDDDDSDGDESEFEPTPKKPASSKRRANETTP 180
Db 121 AVAKPKAKPAEVPKPAVDEDESDSDCMDDDDSDGDESEFEPTPKKPASSKRRANETTP 180
OY 181 KAPVSAKAKAVYVTPKOTDEKKKGAANOSPKSASQVSCGCKTTFNSGNALSHNKAK 240
Db 181 KAPVSAKAKAVYVTPKOTDEKKKGAANOSPKSASQVSCGCKTTFNSGNALSHNKAK 240
OY 241 HAAAK 244
Db 241 HAAAK 244

```

```

RESULT 3
ID 09FNJ6 PRELIMINARY; PRT: 306 AA.
AC 09FNJ6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HISTONE DEACETYLASE-LIKE PROTEIN.
GN ATG22650 OR MDJ22.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneo T., Asamizu E., Miyajima N.,
RA Tabata S.;
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shino P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB006699; BAB1671.1;
DR EMBL: AV059893; AAL24375.1;
SQ SEQUENCE 306 AA; 32348 MW; A19274D43BD0142C CRC64;

```

Query Match 47.0%; Score 600.5; DB 10; Length 306;
 Best Local Similarity 50.5%; Pred. No. 1.1e-37;
 Matches 151; Conservative 24; Mismatches 65; Indels 59; Gaps 14;

```

OY 1 MEFWGLGVKSGKPYVTPREBGLIHVSQASLGECNNKKGFPVPLHVKVGNOMLVLTIST 60
Db 1 MEFWGLGVKSGKPYVTPREBGLIHVSQASLGECNNKKGFPVPLHVKVGNOMLVLTIST 60
OY 61 ENIPOLCDLVDFKPEFLSHTWKGSYFVGYKTPNIEPOGYSEEEEPKPKKRAKETTP 115
Db 61 ENIPOLCDLVDFKPEFLSHTWKGSYFVGYKTPNIEPOGYSEEEEPKPKKRAKETTP 115
OY 116 -----GNAKAV-----AKPKAKPAEVPKPAVDEDESDSD-----GM--D 149
Db 120 AVTANAGAAVYKADTKPKKPAEVPKPAEVPKPAEVPKPAEVPKPAEVPKPAEVPKPAEVP 179
OY 150 EDDSDG---EDS---EEEPKPKKPAEVPKPAEVPKPAEVPKPAEVPKPAEVPKPAEVPKPAEVP 199
Db 180 EDDSDGDEDESDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 239
OY 200 EKKKGKGA-----ANOSPKSASQVSCG--CKTTFNSGNAL--ESHKAK 241
Db 240 EKKKGKGTATTPPKKKGKSPYANANOSPKSASQVSCGSGNNKKNKPFNSGKFGGNNKGN 298

RESULT 4
049209 PRELIMINARY; PRT: 305 AA.
AC 049209;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN HD2 OR HD2B.
OS Arabidopsis thaliana (Mouse-ear cress).

```


DR InterPro: IPR000822; znf_C2H2.
 DR SMART: SM00355; znf_C2H2.1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
 DR DNA-binding: Metal-binding; zinc-finger.
 KW SEQUENCE 294 AA; 31830 MW; 7B69F3BE8324C1D1 CRC64;

Query Match 40.1%; Score 513.5; DB 10; Length 294;
 Best Local Similarity 40.5%; Pred. No. 3,6e-31;
 Matches 121; Conservative 33; Mismatches 84; Indels 61; Gaps 8;

QY 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 60
 DB 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 60
 QY 61 ENIPOLFCDLVFPKFEELSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 120
 DB 61 EKFPQSTELVLRNPLSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 115
 QY 121 AVAKPKAKPA-----EVKPAVDDEDESDDSCMDDEDS----- 153
 DB 116 AAKPSAAKQVNPOLPNEVDYAKAKODDADGSEEDSDSDSNGDEEEKVTAESDSEED 175
 QY 154 -----DGESEEEPTPKPKPASKKRANE-TTPKAPYSAKKAKAVTPQKTDEKK----- 203
 DB 176 DSSDDEEDDSSEETPKPKPEEPKRSAPNSKNPAKAKF-VTPQKTDSKPPHYHVA 234
 QY 204 -----GKAA-----NOSPASAOVSCGCKTKTPNSGNALESHNAKHAHA 244
 DB 235 TPIPSKQAGKNSGGSGTSGTSGKQOTPKSAGAFGCKSCTRTFTSEMGLOSHTKAKISAA 293

RESULT 7
 Q9M4T5 PRELIMINARY: PRT: 297 AA.
 AC Q9M4T5:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PUTATIVE HISTONE DEACETYLASE HD2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocarotidae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN 11
 RA SEQUENCE FROM N.A.
 RA Dangi M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
 RT "Molecular characterization of type-2 histone deacetylases in higher
 plants";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF255711; AAF70196.1; -
 DR InterPro: IPR000822; znf_C2H2.
 DR SMART: SM00355; znf_C2H2.1;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
 KW DNA-binding; Metal-binding; zinc-finger.
 SO SEQUENCE 297 AA; 32502 MW; EEB3AACODZC41C63 CRC64;

Query Match 40.0%; Score 511.5; DB 10; Length 297;
 Best Local Similarity 42.3%; Pred. No. 5.1e-31;
 Matches 126; Conservative 36; Mismatches 79; Indels 57; Gaps 10;

QY 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 60
 DB 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 59
 QY 61 ENIPOLFCDLVFPKFEELSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 113
 DB 60 DKFPQLODFVFPKFEELSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 119

QY 114 PA-----GNAKAVAKPKAPAEKPAVDDEDES-----DG 147
 DB 120 PARANGKEVEKENEKQKGTDSASKAAYVDDDDDAEDSDDEDLSPEEDDD 179
 QY 148 MHRDUS--DGESEEEPTPKPKPASKKRANE-TTPKAPYSAKKAKAVTP--QKTDEKK- 202
 DB 180 SSEDSEDEDEDEDEPTPKPKPEEPKRSAPNSKNPAKAKF-VTPQKTDSK 238
 QY 203 -----KGGKAANOSPASAOVSCGCKTKTPNSGNALESHNAKHAHA 243
 DB 239 VHAVTPHAKAKSKTPVNDKSKPEKSPKSGCSISCKSKTPTNSMALOSHAKHAPA 296

RESULT 8
 Q9LZR5 PRELIMINARY: PRT: 296 AA.
 AC Q9LZR5:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE HISTONE DEACETYLASE-LIKE PROTEIN.
 GN F17C15_160.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN 11
 RA SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 RN 12
 RA SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL162506; CAB82939.1; -
 DR InterPro: IPR00822; znf_C2H2.
 DR SMART: SM00355; znf_C2H2.1;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
 KW DNA-binding; Metal-binding; zinc-finger.
 SO SEQUENCE 296 AA; 32129 MW; 4B2DABEALFIE9D CRC64;

Query Match 39.0%; Score 498.5; DB 10; Length 296;
 Best Local Similarity 39.5%; Pred. No. 4.9e-30;
 Matches 121; Conservative 31; Mismatches 81; Indels 73; Gaps 9;

QY 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 60
 DB 1 MEMFGIEVKGSKPVYTPREGILIHVSQALGECNNKGEFVPLHYKVGNONLVLTIST 60
 QY 61 ENIPOLFCDLVFPKFEELSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 120
 DB 61 EKFPQSTELVLRNPLSHTWGKGSVYFVGYKTPIEPGVSKEEEEEEVPAAGNAK 110
 QY 121 -----AVAKPKAKPA-----EVKPAVDDEDESDDSCMDDEDS----- 153
 DB 111 WGLVTPAPAKSAKQVNPOLPNEVDYAKAKODDADGSEEDSDSDSNGDEEEKVTA 170
 QY 154 -----DGESEEEPTPKPKPASKKRANE-TTPKAPYSAKKAKAVTPQKTDEK 201
 DB 171 ESDSEEDSDDEEDDSSEETPKPKPEEPKRSAPNSKNPAKAKF-VTPQKTDSK 229
 QY 202 KR-----GKAA-----NOSPASAOVSCGCKTKTPNSGNALESHNAK 238
 DB 230 KPHVAVATPHPKQAGKNSGGSGTSGTSGKQOTPKSAGAFGCKSCTRTFTSEMGLOSHTK 289
 QY 239 AKHAAA 244
 DB 290 AKHAAA 295

RESULT	9	
ID	09M4U4	
AC	09M4U4	PRELIMINARY; PRT; 300 AA.
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	HISTONE DEACETYLASE 2 ISOFORM C.	
OS	zebra mays (Maize).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;	
OC	Panicoidae; Andropogoneae; Zea.	
OX	NCBI_TaxID=4577;	
RN	11	
RP	SEQUENCE FROM N.A.	
RA	Dargyl M., Brosch G., Haas H., Loidl P., Lusser A.;	
RT	"Molecular characterization of type-2 histone deacetylases in higher	
RT	plants.";	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF254073; AAF68625.1; -	
DR	InterPro; IPR000822; Znf_C2H2.	
DR	SMART; SM00335; Znf_C2H2; 1.	
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.	
DR	PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.	
KW	DNA-binding; Zinc-finger.	
SO	SEQUENCE 300 AA; 32457 MW; 62FECA148B1E1FE CRC64;	

Query Match	38.5%	Score 493	DB 10	Length 300
Best Local Similarity	42.1%	Pred. No. 1.3e-29		
Matches 128	Conservative 35	Mismatches 71	Indels 70	Gaps 13

QY	1	MEMGLEVKSCKPVTATVPEEGCILLIHVSQASUGSCNKKKGEFVPLHAKVGNOMLVIGTST	60
Db	1	MEMGLEVKSCKPSTATKCEPGGFGVLHLSQALGE--SKSDNMLMYKIDDKALATGTL	58
QY	61	ENIQLQCDLVDEDEFEFLSTHWKGSVYFVGATPTNIEPQSYSE-----KEEEEE	113
Db	59	DKNPIDFLIFDKEFELSHSTKSTTSYFFGVAV---RQPEDEBMDLDSEDEBELNV	114
QY	114	PA-----GNA-----AKAVANP-KAKPAEVPAYVDEDESDSGMDESDS	154
Db	115	PAVENGKADKKQKQSGEKAVAAPSKSSPDSKSDDDSEDEDETDSEDETDSDSEGL	174
QY	155	-----GEDSEEEPTPKPKPSSKKRANETTP-KAPYSAKKAAVATP--Q	196
Db	175	SPEEGDDSDSDDEDDTSDDEEDPTPKPKPEVGGKRAAESVLTPLPSDKKAVVA--TPSSQ	233
QY	197	KTDKE-----KGGKAN-----QSPKSA--SOVSGCGCKTGFNSGNALESH	237
Db	234	KTGCKKGAAVATPHPRAKCKTIVNNDKSKRSKPSAKPSVPPCKSCSKSFISETAPQAH	293
QY	238	KAKH 241	
Db	294	KAKH 297	
RESULT	10		
OZ4591			
AC	OZ4591	PRELIMINARY;	PRT; 307 AA.
DT	01-JAN-1998	(TREMblurel. 05, Created)	
DT	01-JAN-1998	(TREMblurel. 05, Last sequence update)	
DT	01-DEC-2001	(TREMblurel. 19, Last annotation update)	
DE	NUCLEOLAR HISTONE DEACETYLASE HD2-P39.		
GN	HD2.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade		
OC	Paricoidae; Andropogoneae; Zea.		
OK	NCBI_TaxID=4577;		
RN	[1]		

RP SEQUENCE FROM N.A.
 RC STRAIN-CUZCO 251;
 RX MEDLINE=97349336; PubMed=9204905;
 RA Lusser A., Brosch G., Loidl A., Haas H., Loidl P.;
 RT "Identification of maize histone deacetylase Hb2 as an acidic
 RL Science 277:88-91(1997).
 RM [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CUZCO;
 RA Brosch G., Lusser A., Loidl P.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBB databases.
 DR EMBL: U82815; AAB63262.1; -;
 DR EMBL: AF026917; AAC61674.1; -;
 DR InterPro: IPR000822; Znf-C2H2.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 KW DNA-binding; Zinc-finger
 SO SEQUENCE 307 AA; 35238 MW; 9CFF6B0360274FC1 CRC64;

Query Match	37.7%	Score 482	DB 10	Length 307
Best Local Similarity	39.0%	Pred. No. 8	Re-29	
Matches 121	Conservative	38	Mismatches 79	Indels 72
				Gaps 11

OY		MEMFQIENSKRPATVTPPEEGILLIHVSQAQSIECKNNKKGEEVPLHVVGNQLVLGTST	60
Db		1 MEMFQIENSKRPATVTPPEEGILLIHVSQAQSIECKNNKKGEEVPLHVVGNQLVLGTST	60
		1 MEMFQIENSKRPATVTPPEEGILLIHVSQAQSIECKNNKKGEEVPLHVVGNQLVLGTST	60
OY		61 ENIPOLFCDLVFDKEFEELISHTWNGSVYEVGCTKPNIIEPCGYSE-----EEEEEEEV	113
Db		59 DKNPHIQDILLIFDEEFELSHTSKTTSVFETGYKV----EQPFEEDEMOLDSEDELELV	114
OY		114 PA-----GAA-----AKAAYP-KAKPAVKRAYVDDEEDSESODMDDDSD-----	154
Db		115 PAVKENCAIDKKOKSQPKATAAPASKSPPDSKKDDDDSDDEDDETDSDEDETDOSDECL	174
OY		155 -----GEDSEEEEPPTPKKPASSKR-ANETTPKAPVASAKKAYAVTP--Q	186
Db		175 SSEEDDDDSDDEDJTDDEEDDTPTPKPEVGKRRPRESSYLTKPTLIDKKAKAVA--TPSSQ	233
OY		197 KTDEKK-----KCGRAANOQSKSAOVSQCSGCKRTFPNGNALES	235
Db		234 KITGGKGAAYVAHPHPAKKTIYNNDKSVKSPKSAFKSGGSVPCKPCSFISTETALQA	293
OY		236 HNKAKHAAPK 245	
Db		294 HSRAMGASE 303	
RESULT	11		
ID	O94f81	PREDIMINARY; PRF;	286 AA.
AC	O94f81;		
DT	01-DEC-2001 (TREMBLrel_19, Created)		
DT	01-DEC-2001 (TREMBLrel_19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel_19, Last annotation update)		
DE	HDI2 TYPE HISTONE DEACETYLASE HDAl06.		
GN	HDAl06.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade		
OC	Panicoidae; Andropogoneae; Zea.		
OX	NCBI_TaxID=4577;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN-CV. B73;		
RA	Chandler V.L., Kaeppler S.M., Kaeppler H.F., Cone K.C.;		
RT	"Sequences from the Plant Chromatin Consortium (NSF Plant Genome		
RL	Project 9975930)."		
RM	Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.		
RM	(2)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CV. B73;		

OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96019267; PubMed=7593294;
 RA Cairns C., McStay B.;
 RT "Identification and cDNA cloning of a Xenopus nucleolar
 RT phosphoprotein, xNopp180, that is the homolog of the rat nucleolar
 RT protein Nopp140.";
 RL J. Cell Sci. 108:3339-3347(1995).
 DR EMBL: X88927; CAB61368.1;
 SQ SEQUENCE 990 AA; 102025 MW; E62054C86A3B4713 CRC64;

Query Match 12.9%; Score 165.5; DB 13; Length 990;
 Best Local Similarity 30.7%; Pred. No. 0.0002;
 Matches 55; Conservative 17; Mismatches 64; Indels 43; Gaps 7;
 QY 103 SEEEEEEEVPAGN-AAKAVAKPKAPAEVYPVADDEDESDSDGDEDDSD-----154
 DB 703 SSESEEGSKOPTGKSPAKATAPPKNPVAVNKDPSSSSSSSDSGDDEKQKPKQAAA 762
 QY 155 ---GDSSEEEPTPKKPKASK-----KRANE-----TPPKAPVSAKKAKV 191
 DB 763 KDVKQGAAGAATPKPKAASSESDSDSDVSKAKKTAVSKSPVTPPKAVPAKAKSS 822
 QY 192 AVTPQKTEKKKG-----GKAANSPKASGVSCGCKKTFSGNALSHNKAKHAAAK 245
 DB 823 SSSSEDEKQCKNTSTTKIANSTPKAAA-----AECSESSSSS---EDEGRANGTSGK 874

RESULT 15

Q9STR2 PRELIMINARY: PRT: 487 AA.

AC Q9STR2:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 53.3 KDA PROTEIN.
 GN T30C3.20 OR AT4G25340.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL079350; CAB45512.1;
 DR EMBL: AL161563; CAB81345.1;
 DR HSP: Q00688; 1PRK.
 DR InterPro: IPR001179; FKBP_PPase.
 DR Pfam: PF00254; FKBP_1.
 DR PROSITE: PS00454; FKBP_PPase_2; 1.
 DR PROSITE: PS50059; FKBP_PPase_3; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 487 AA; 53290 MW; 5B2FA21570BC0AC6 CRC64;

Query Match 12.8%; Score 164; DB 10; Length 487;
 Best Local Similarity 26.5%; Pred. No. 0.00012;
 Matches 83; Conservative 33; Mismatches 87; Indels 110; Gaps 16;
 QY 1 MEFW-----GIEKSKRPVTVP--EEGLIHVSQASLSECKNKGEVPLHVKV 48
 DB 1 MGFMSDFVLLTVAGLEKPKPKQAYNPKNKQG-KIHVQATLTGTGSKKSVI--OCST 57
 QY 49 GNO-NLVLTSTENIPOLFCDLVDEKFEFELSHMGSGYFVGYSKTPNIEPGYSEEE 107
 DB 58 GDKAPIALCSLIPNKIECCPLNLEFDDDE-----PVEFTVGRSHLSGFLLEYQ 109
 QY 108 EEEVEVPAGNAKAVAKPKAPAEVYPVADDEDESDSDGMD-----EDSDGDSSEEE 162
 DB 110 DDED-----DYEHDESDSDGIDVSESEDESDSCYDSEDE 144
 QY 163 -----PTPK-----KPASSKRNETTPAPVS---AK 187
 DB 145 QLDEFEDFLDSNLERYMAAPKSGVILEIEDEKPAKDNK-AKQTKKKSQAEGENAK 203
 QY 188 KAKYAV-----TPQKTEKKKG-----GKAANSPKASGV-----SCGSK---KTF 227
 DB 204 KQYVIEGAHVPLESEDEDEGLPIPKGKSSEVENASGEMVAVDNDGSKNRKAKAA 263
 QY 228 NSGNALSHNRKAK 240
 DB 264 EQDDQGESANKSK 276

Search completed: May 2, 2002, 22:30:22
 Job time: 1583 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 20:13:40 ; Search time 5336.92 Seconds
(without alignments)
4775.886 Million cell updates/sec

Title: US-09-645-337-7
1218
Sequence: 1 gcttcgcgtcctcaaaaaa.....aaaaaaaaagcgcgccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_la:*
19: em_mu:*
20: em_or:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hgc_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1218	100.0	1218	6	AX139020	AX139020 Sequence
2	1116.4	91.7	1172	8	AF195546	AF195546 Arabidops
3	1099.4	90.3	1187	8	AY059893	AY059893 Arabidops
4	847.4	69.6	918	8	AF044914	AF044914 Arabidops
5	309.4	25.4	77363	8	AB006699	AB006699 Arabidops
6	159.2	13.1	977	8	AY065396	AY065396 Arabidops
7	156.2	12.8	914	8	AF195545	AF195545 Arabidops
8	156.2	12.8	939	6	AX139018	AX139018 Sequence
9	146.4	12.0	994	8	AF255711	AF255711 Oryza sat
10	112.8	9.3	903	8	AF254073	AF254073 Zea mays
11	112	9.2	1307	6	ARI68371	ARI68371 Sequence
12	111.6	9.2	1191	6	ARI68369	ARI68369 Sequence
13	111.2	9.1	1245	6	ARI68370	ARI68370 Sequence
14	111.2	9.1	1283	6	ARI68368	ARI68368 Sequence
15	110	9.0	1169	8	ZMU82815	U82815 Zea mays nu
16	107.4	8.8	912	8	AF254072	AF254072 Zea mays
17	100.6	8.3	1124	8	AF384033	AF384033 Zea mays
18	95	7.8	101371	8	AC002534	AC002534 Arabidops
19	90.4	7.4	1236	8	AF372889	AF372889 Arabidops
20	85.4	7.0	7218	6	166494	166494 Sequence 14
21	84.4	6.9	885	8	AF255712	AF255712 Arabidops
22	68.6	5.6	2486	8	AF026917	AF026917 Zea mays
23	68.4	5.6	1397	10	BC003489	BC003489 Mus muscu
24	68.4	5.6	1408	10	BC005628	BC005628 Mus muscu
25	66.8	5.5	1127	10	AB025582	AB025582 Mus muscu
26	65	5.3	209213	2	AL611930	AL611930 Mus muscu
27	63.8	5.2	194731	2	AC079365	AC079365 Mus muscu
28	62	5.1	132323	8	AP003416	AP003416 Oryza sat
29	62	5.1	174218	2	AC091071	AC091071 Oryza sat
30	61.6	5.1	170676	2	AC023448	AC023448 Homo sapi
31	60.8	5.0	75475	8	AB005235	AB005235 Arabidops
32	60.8	5.0	102897	8	ATF17C15	AL162506 Arabidops
33	60.8	5.0	201237	2	AL606932	AL606932 Mus muscu
34	60.6	5.0	55296	2	AC101779	AC101779 Mus muscu
35	60.4	5.0	70266	2	AC102233	AC102233 Mus muscu
36	60.4	5.0	205402	2	AL645584	AL645584 Mus muscu
37	60	4.9	181238	2	AC094176	AC094176 Rattus no
38	60	4.9	203942	2	AL671331	AL671331 Mus muscu
39	59.8	4.9	73132	9	HS83382	AI023800 Homo sapi
40	59.6	4.9	178820	9	AC020552	AC020552 Homo sapi
41	59.6	4.9	181230	2	AL358876	AL358876 Homo sapi
42	59.4	4.9	204783	2	AL645857	AL645857 Mus muscu
43	59.2	4.9	195028	30	AC024739	AC024739 Homo sapi
44	59	4.8	139665	2	AP000865	AP000865 Homo sapi
45	59	4.8	205638	2	AC040927	AC040927 Mus muscu

ALIGNMENTS

RESULT 1
AX139020 1218 bp DNA linear PAT 30-MAY-2001
LOCUS AX139020 Sequence 7 from Patent EP1094112.
DEFINITION AX139020
ACCESSION AX139020
VERSION AX139020.1 GI:14274701
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.

REFERENCE
AUTHORS Wu, K., Miki, B.L., Tian, L. and Brown, D.C.
TITLE Repressing gene expression in plants
JOURNAL Patent: EP 1094112-A 7 25-APR-2001;
The Minister of Agriculture and Agri-Food (CA)
FEATURES
source 1..1218
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 395 a 237 c 296 g 290 t
ORIGIN


```

Db 121 GTGGTTTGGAGTGTGACTGTTGGTGGGCTAAACTGTTATGCAACACTTTCACAAAGAC 180
Oy 241 aaattccctcagattagcttgaattggttttgcataaagaattgaacttccacacac 300
Db 181 AAGTCCCTCAGATTAGCTTTGATTGGTTTGTGTTAAAGAGTTTGAGCTTTCACACAGC 240
Oy 301 ggtaccacaaagaatgttcaatttccattggtcacaatcccccaaatcgaagatgac 360
Db 241 GGTACCAAGCAAAATGTTCAATTCATTGGTACAAATCCCCCAACATCAGACAGATGAC 300
Oy 361 ttcactaattcggatgataaggaatttctcgaagctgttctgcctccctccactgct 420
Db 301 TTCCTACTGTTGGATGATGAGATGTTCTTAAGCTGTTCTGCTCCCTCCCTACTGCT 360
Oy 421 gttactgcacacgaaatgtctggaacagctgttgcagagctgcacacaaagcgaagcc 480
Db 361 GTTACTGCCAAGGAAATGCTGGAGCAGCTGTTCGCAAGGCTGACACAAAGCAGAGCC 420
Oy 481 aaacctgcgaagtgaaagcctgcgaagaagaagcctgaatcagaaggaagatgactct 540
Db 421 AAACCTGCCAAGTGAAGCTGTCAGAGAGAGAGCTGAAATCAGACGAGAGATGAGTCT 480
Oy 541 gatgtgaaagatgaaagctgaaagaagatgatactctgaaagaagaagatgattgat 600
Db 481 CATGTGAGAGATG-AGTCTGAAGAGGATGATGACTC-TGAGAAAGGAGATGAT-GTTGAT 537
Oy 601 gaagatgactcagatgataagagaagaagatctgaagatgaaagaagaagagag-c 659
Db 538 GAAGATGACTCAGATGATGACGAGAGAGG-AGGATTTCTGAGATGAAAGAGAGAGAGAGC 596
Oy 660 tcttaagaagcctgagcacaatcacaagaagaagcacaatgaaatcgtatcccaaacac 719
Db 597 TCCTAAGAAAGCTGAGCAATCAACAGAGAGGCAATGATGATATCCAAAGACAC 656
Oy 720 cgtctctgaaagaagagcaaacacacagcagcagcagcagcagcagcagcagcagc 779
Db 657 CCTCTCTGAAAGAGAGCAAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 713
Oy 780 aagaagaagaagaagaagacacacacacacacacacacacacacacacacacacacac 839
Db 714 AGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
Oy 840 tccctgtgaatgctaacagagcccaagctggaagtgatcatcatcctcgtgtgtaacaa 899
Db 774 TCCTGTGAATGCTAACAGAGAGCCCAAGCTGAGAGTCAATCATCCGGTGTGACACAA 833
Oy 900 caagaagcattcaactcagcgaacaaatgttggtgttcccaacaaaggttctcaaca 959
Db 834 CAAGAAGCCATTCAACTCAGCCAAACAAATTTGTTGTTCCAAACAAAGGTTTCTTAACA 893
Oy 960 gggcaaggaagaggtagagcttaa 984
Db 894 GGGCAAGGAGAAAGGTAGAGCTTAA 918

RESULT 5
LOCUS AB006699 77363 bp DNA linear PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MDJ22.
ACCESSION AB006699 BA000015
VERSION AB006699.1 GI:2351064
KEYWORDS
SOURCE
  Arabidopsis thaliana (strain:Columbia) DNA, clone_lid:Mitsui P1
  clone:MDJ22.
ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
  1 (sites)
  Koriaki H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.
  and Tabata S.
  Structural analysis of Arabidopsis thaliana chromosome 5. II.
  Sequence features of the regions of 1,044,062 bp covered by

```

```

JOURNAL
MEDLINE
  DNA Res. 4 (4), 291-300 (1997)
REFERENCE
  2 (bases 1 to 77363)
AUTHORS
  Nakamura, Y.
TITLE
  JOURNAL
  Direct Submission
  Submitted (22-AUG-1997) Yasukazu Nakamura, Kazusa DNA Research
  Institute, Department of Plant Gene Research, 1532-3, Yana,
  Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamukazusa.or.jp,
  Tel:81-438-52-3935, Fax:81-438-52-3934)
  Address for correspondence: kaos@kazusa.or.jp
  For the latest information on annotation of this clone, please see
  http://www.kazusa.or.jp/kaos/cgi-bin/agd.graph.cgi?c=MDJ22
  Genes with similarity to proteins in the databases are described in
  'product' or 'note' qualifiers. Genes that have no significant
  protein similarity are described as 'unknown protein'.
  The software programs used to predict genes include: Grail
  (Informatics Group, Oak Ridge National Laboratory,
  http://compbio.ornl.gov/Grail-1.3/),
  GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
  NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
  Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
  SplicePredictor (Volker Brendel, Stanford University,
  http://gremli.zool.jastate.edu/cgi-bin/sp.cgi).
  Genes encoding tRNAs are predicted by tRNAscan-SE
  (Sean Eddy, Washington University School of Medicine, St. Louis,
  http://genome.wustl.edu/eddy/tRNAscan-SE/).
  This sequence may not be the entire insert of this clone. It may be
  shorter because we remove overlaps between neighboring submissions.
  The 5' clone is MDJ16 and the 3' clone is K5A21.
  Location/Qualifiers
    1..77363
      /organism="Arabidopsis thaliana"
      /strain="Columbia"
      /db_xref="taxon:3702"
      /chromosome="5"
      /clone="MDJ22"
      /clone_lid="Mitsui P1"
      /join(1014..1043,1182..1414,1564..2029,2131..2316)
      /note="gene_id:MDJ22.1"
      unknown protein
      /codon_start=1
      /evidence=not_experimental
      /protein_id="BA011665.1"
      /db_xref="GI:10178233"
      /translation="MGPYLDGPKKILLYLPTKRAVSVLSKRMKSLMSLPGLDFDSE
      FPNATFVSMDRPFVDSREKSKLHLKLTISKVMSDQRTLRSLISLPRKTVHL
      OLNVYANDAGLIESFISYPAREDFSIYRSVDIDNNVNVVNVHSDTLTSLSEVDFLGEW
      GLGILCGOAGLGLMIDDPFRKFLNFHEHLSDEKTLISNLVYKVPVGVGYNHSGYGL
      SKOOMARNVFTGISRYRDLIISADIFDICYLKEESLDPFCNLISCLAEVGVSSVWK.IRE
      TLQKPEPCNKLSKSLFLVMTCTNPRALT"
      /join(6501..6675,6836..7458,7550..7699,7802..8110)
      /note="gene_id:MDJ22.2"
      unknown protein
      /codon_start=1
      /evidence=not_experimental
      /protein_id="BA011666.1"
      /db_xref="GI:10178234"
      /translation="MEAREKERICLEPAREDLSKLPNALISHIRIRGQRLSGLY
      FYPGKGVSGWPEVDYVARGNQLHDLDELGCYVFRYVWMPVSLYICTLILKLY
      RVSGNPFESVSLPHLKTLCLEARNLIPNEASVESLISCPVLEDTLTLIRINDKVLRY
      HSGLTFSVGYHGGDLIRKYNYSKVRNSGLVIDAPRLKYLTFNYERSKSKIKRN
      LSLVYKFNSSPDISSVAGCSQGMHNFCTISRYRDLITISRDMMRLTFEYLKNS
      LPFCNLSTYKTLNLSERSLAFDLITFLTLESCTRMKSLALGLCTFLDDMKSVKSV
      KCLDLSLEFVEIKNEHPDGVGLVLAFLFVENSVNKLKLLHFFFLBGNPAVLNDL
      LALPRSSMCOIEFVNPQSGRSRGIF"
      /join(11695..12657,12751..12900,12999..13394)
      /note="gene_id:MDJ22.3"
      unknown protein
      /codon_start=1
      /evidence=not_experimental
      /protein_id="BA011667.1"
      /db_xref="GI:10178235"

```

```

/translacion="MEPIEKRIYVBOIAIEDLISKLPEVILSOILSYLPTKDIYRTS
VLSKHWKSWMLIPGIDIDSSPEPHYDTPVPMNIEPLPSREBNCLKIKKISIOKIK
NDPSCYTLTDCVARKGLOHLDVERGRMEREPEDMPLSLYICKTLHLHLYVL
GNFDOSVDLPSLKSCLTEENYVSNESLSSISGRVLEDTIVKIDNVEFLVHS
QSLTSLSVGKSYGPEIYVYDRDNGSLVADPRLKYLTFNNDOSKISINSGS
LKVYTIICGPKISRYVCGTEQOMAKFLTGISRAVLIYSEDMVEVYSKYLEDSPQ
PCLNLTAKASWMSFDFDLITPKLLESCPNKISIVLETTCTVDRKATVEERASVP
ECLLSLSEFEIKNRTSYDDGALVARYEVENSVOYVILRIASSPIARCNOAALKD
IIELEPRSSMCOIEFNALNGHALCFRRKKTITGRVFLDFVDAESTIGATYS"
complement(join(14209..14547,14624..14731,14851..14949,
15047..15112,15289..15381,15541..15659,15755..15971,
16065..16170,16400..16701))
/notes="dbj|BAA92923.1
gene_id:MDJ22.4
similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB11668.1"
/db_xref="GI:10178236"
/translacion="MISLPITPTIPSRCLIAKTRONSTRRLIRSSSSIQDDPTV
ETTKRYVLRHGOSTWNEGRIGQSSDSVLTAKGSOAEISROMLIDSEFVCTSP
LRSKRTAETIWSRESEMIEDYDLREIDLYSFOGLKKEGKEGEAFKQMEDPAN
FIIDGHPVRELMSRSCWPGILAHESKSVLVAHNNAQALLATLIGLTERSL
IOSNCGVSLDIPRADGSPHVCNLRLNOTPNPLAGSSGGRKASKOILVCHGOG
NMEDSAVIMQANNDQAMNMLGVISOKTAEILLDRYSSIVCSPTKASIESGVISR
VOBAGCICVDNVPRTVKTOMNELDVSVLSKSKNDVYIASQIDEBFSLAMRSE
KAMESLDELSDKSNPCEIYVVGPMAMHISLIAQCLNTEALGLFHLGDSISVI
DEPDGSSKGVIRCTNYTHLRMSIPITKPA"
21123..22400
/notes="gene_id:MDJ22.5"
/codon_start=1
/evidence-not_experimental
/product="cholesterol mutase/prephenate dehydratase-like
protein"
/protein_id="BAB11669.1"
/db_xref="GI:10178237"
/translacion="MOTISAPSCDLKSYIOPNLTAKKARSVHNGKRVSCYRSE
SFSPNGVSSRADMOSSCALLASKVSKSNSSVAVNNGHNSGVDSLVSCKOHN
GKGLIOPITLITDLPSPHSGSTLVAIVGVGATSEAAAGATVPNSEITPCDDQVA
FOAVELMIDRAVLAQTEVNSLIGSIHNRNDLLRHRLHYGEVOIVHCLALPGVRT
DCITTRVISHDPAQTEVNSLIGSIHNRNDLLRHRLHYGEVOIVHCLALPGVRT
ALVYGLIADIODAGNVTREFLMARPILIPTRDPRKTSIVFAAOPHKTSTVLFK
VLSAFAFNIISLTKIESRPHONCPVRYVDENVGTSKHETFFYVDFEASMAEAAON
ALAVOETSPILRYGSPMDKTPMSTLPSERV"
join(25901..26553,26964..27035,27152..27241,27328..27501,
27723..27813,27972..28034,28129..28408,28466..28569,
28648..28774,28862..29628,29712..30119)
/notes="gene_id:MDJ22.6
unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB11670.1"
/db_xref="GI:10178238"
/translacion="MANEELTESQOQEDPSOOLPMANDEKSDSDNSDSDASQSG
DIPVISESENEARSGNTIFNVYRPDIPDPANPENTIRRNRYLDDCKRYRMQBE
EKYTYTERLPPRDPPEKMKQDLREIADPPLMKTPMDPADEMDVDVDEI
QSGKQDQIOPETVYRKPYPALPDNHYDENKGVVEEDRLEELAWVSYTFPGSS
YEGTWDLDLAQGVYIAENGLVYEGEVLQDMGHGYDIDIDIEIPGSKLEAK
MRAEGRILKRDYMPEDRKMLEMDVEDSVALTQNFQYFENEMVYICAHISGLAEVAA
RYAGOMKHSRMHGGGVVENRILYGRYFPELLLEEGGCTEMVYICAHISGLAEVAA
KARMFVNPDSIAPLIYHPCVLVICTYKPRHIGICILIPDOCRITSSNNHSNNT
YRPIISGNVQLMKRYKKTGHFTVRLIREKRYGDPDHPYTFEEDVMAAPQFIMQY
EVPETWETVGEVDODEREMNLSFYKAPRLPMLPELEHMMANVEVPEPEVLNLEPE
PDPDPSKLVOKEDVILHTPGRIINVEDKHOIRLFMOPPLEGGEVDSKTEFL
PQDPEFYGVKVVVKKHPIKSFVLIEGIEVSPMLDGLKTEEMKKEVYEEKKEIOO
ELIEVAELICLEALIDMOEELKKOEKKEKTEMGILFEEDDVAIVPYKKEVYTAK
EKIOENKOEKYYKDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
FASCTIPAVOSRLSSSLAMKOHARPSKVNTGILKGDVTSASITHPPLPSNARL
KMGKVAANRCGVORSYSSRSQSLMSLSLSSCNSSSSSPDSSSEYLLKQSGIMET
PVGDSVYSLDIOITKCDLFAETPAVS"
join(30598..30610,30871..30941,31032..31240,31371..31450,
31536..31667,31802..31882,31983..32091,32233..32359,
32454..32532)

```

```

/notes="gene_id:MDJ22.7"
/codon_start=1
/evidence-not_experimental
/product="histone deacetylase-like protein"
/protein_id="BAB11671.1"
/db_xref="GI:10178239"
/translacion="MEKGVATVTKNATKVPPEBDSLVHISQASLDCTVSGSEVYLS
VTVGAKVIGITLSODKRPQISFLVPEKEELSHSGTANVHTIGKSPNIEQDDEFT
SSDDEVEAEPAPAPVATVANGNAAVVAADKPAKPAVEVPAEKPEPESDEDS
DDEDESEDDVSGEMVDEDDSDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
PVSCKAKPAAAPASTPQTEKKGKGTATPHAKKAGKSPVANOSPCKSGQSSGCG
NNKKPFPNSGPOGSSNNKSGNKGKGRARA"
complement(join(32806..33093,33104..33324,33408..34331))
/notes="gene_id:MDJ22.8
unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB11672.1"
/db_xref="GI:10178240"
/translacion="MSNOGAIKSGENDRISLPHLLSOLLISNLPTENAVTTSLSTR
WKDLMLSTPVLDIDIDADADATTFISFAITRFLDSFKSLKLIQISOMEAVDMKTI
PMIEDAVVRRRIQHLVEDSRIDMIDTLPITYLESVSLRLHLMHREFVSLPL
KVMHLENTYVETMEREFISSCVLEDTIVRVRVDEATEKVLVSSOSLNLKLYID
SSKCYNDSDMDKVVYIDAPOLYVLTIKDDOSFPVFNINLCSKAKKADIKSPFNSDIH
DLSEFERNKPELTGSSLRDMITISCTTLKTIKCHUKHEPMQFPMNTRLHAKTV
CDLEMLPCVLSKCPVNLKSLVLIKLGEMENRISLSSVVPKQJSSLACEIIRPVYS
GEEMLSYFLENISLVLEKFKLCRCHSEBESLVREIMTFORCSSCEINVRFOR"
complement(join(34670..34948,35032..35172,35252..36163))
/notes="gene_id:MDJ22.9
unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB11673.1"
/db_xref="GI:10178241"
/translacion="MSTRKGYTSQDSISLIPDULCRILSNLPTKVAVRYSKRRK
RPSISVPLENVSERFQYTFEPAVYHGFIDTSEPTCHIKLIKLAPEKKOHRSTLTM
IHMVKRVHOLDIGRMVSYLQDELIPHSLYTCETLVSLRLHNSVLPDFDHVSLPRLKT"
25.4%: Score 309.4; DB 8; Length 77363;
Best Local Similarity 81.3%; Pred. No. 3,8e-62;
Matches 414; Conservative 0; Mismatches 1; Indels 94; Gaps 1;

Query Match
775 acagaagaagaagaagagacacacgcacacacacacagcgaagaagggtgga 834
|||||
32229 ACAGAAAGAAAGAAAGAGAGACACACGCCACACACAGCTAAGAAAGGTGGA 32288
|||||

835 aggtccctgtaatgctacacgaagcccaagcttggaagctaatccgtgtgaac 894
|||||
32289 AAGTCTCCTGTAAAGCTAACACGACCCCAAGTGTGAGGTCATCCGCTGTAAAC 32348
|||||

895 aacaacaaga----- 905
|||||

32349 AACAAACAACAACTGCTCTCTCTTTTATCTTTTCATAGCCCAAGATATAGTGA 32408
|||||

906 -----gcattcaactcagg 920
|||||

32409 AATGTTAATGTGTAATGATGTGTAATGTGTGTGTACAGGCAATCACTCAAG 32468
|||||

921 caaacaatttggttccacaacaagaaggtctcaacaaggcgaagggaagggtgaagc 980
|||||
32469 CAACAATTTGTGCTCCAAACAAGAGGTTCTAACAAAGGCGCAAGGAGTAGAGC 32528
|||||

981 ttaagagctggaatcaagaagaggttctgagtttcgaagtagaagataaacaacttga 1040
|||||
32529 TTAAGAGCTGATCAAGAGAGGTTTGGTTTCGACTAGATATCAAAAACATTGGA 32588
|||||

1041 agtgtgatttggaatttatctatattatagataactgttatcgaatgaatcatt 1100
|||||
32589 AGTGTGATTGGAATTTATCTATTATATAGATACTGTTATGAGATGAGACTATT 32648
|||||

1101 ttgagattattgcaattctcattctatgtaattcagtagaataattgctggaatt 1160
|||||

```

CDS

CDS

CDS

CDS

CDS

Db 3269 TTTGAGTATTTGGCAATTTCTACTTTCTATGTAATTCAGTATATGAAATATTTGCTGAAATG 32708

QY 1161 agaaagagagctcgaattgcacaacaaaa 1189
|||||
Db 32709 AGAAGAGAGACTGCAATTCGAAACACAAA 32737

RESULT	6
AY065396	
LOCUS	
DEFINITION	977 bp mRNA linear PLN 11-DEC-2001
ACCESSION	AY065396 Arabidopsis thaliana putative histone deacetylase (At3g44750) mRNA,
VERSION	complete cds.
KEYWORDS	AY065396.1 GI:17529219
SOURCE	FLI.CDNA
ORGANISM	thale cress. <i>thaliana</i>

The Salk, Sanford, PEC (SSP) Consortium members carried out the sequencing and annotation of the RATL cDNAs: Yamada, K., Banh, J., Chang, C. H., Chung, E., Dale, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Toriuni, M., Wu, H. C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shinn, P., Southwick, A., Davis, R. W., Ecker, J. R. and Theologos, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES
source

gene
5'UTR

CDS	3' UTR	BASE COUNT	ORIGIN
/gene="At3g44750"		311 a	183 c 231 g 252 t
63..800			
/gene="At3g44750"			
/codon_start=1			
/evidence=experimental			
/product="putative histone deacetylase"			
/protein_id="AI138837.1"			
/db_xref="GI:17529220"			
/translation="MEFMGIETVSGKRPVTYTPERGILIHVSOASIGCGKKKKEFVPL			
HYKGNQNLVIGTISTENITPOLFCDDLVFDKBEFELSHTWGSGYTFVGYKTPNIEPQST			
EEPEEEEEEEVPAAGNAKAAVAKPAKPAKPAKPAVDDEDESDSGMDSDSDSDXIDESEE			
EEPPPKPASSKRRANETTPAIPVSAKAAVAVYPTDDEKKKGKGAANOSPPASASQV			
SCGSGCKTFFNSGNALESHNAKHAATAAK"			
801..977			
/gene="At3g44750"			

Query Match	13.1%	Score 159.2	DB 8	Length 977
Best Local Similarity	62.4%	Pred. No. 5.4e-27		
Matches 267	Conservative 0	Mismatches 156	Indels 3	Gaps 1

QY	2	tccttcgcttcctaaaaaaaacctaaacacctctctctctctccttcctcgttcacaaacaa	61
Db	4	TTTTAGCTGCTCATAAACCTTAAAAATTCCTCTTTTCTCAACCTTGATTTCTTAGCCA	63
QY	62	tggagctctcgggagcttgcgcgtgacacccaaaaacgcctactaaagtgactctgaagaag	121
Db	64	TGCAGTTCTGGGGAAATTGAAGTTAAATCAGGAAGCCAGTTACAGTGACTCCTGAAAG	123
QY	122	acagcctctgcacactctctagcttcaact---gactgacacgtgaactctggagaat	178
Db	124	GCATTCTTATCCAGCTTCTCAGGATCGCTTGAGATGTAAAAACAAAGAGGAAGT	183
QY	179	ctgtgtcttcagagctgcagcttgctgctggcctaaactgttatctggaacacttccacaa	238
Db	184	TGTGCTCTTTACATGATAAGGTGGGAACGAAAGCTTGCTGGAAGCTATACGACTG	243
QY	239	acaagctccctcagaatagctcttgacttgcttcttgatlaaagaagcttgaaccttcacaa	298
Db	244	ACAAATCCCTCAGCTTTCTGTGATTTGGTATTCGCAAGAAAGATTGGACCTTTCTACA	303
QY	299	ggcgtcaccaagcaacatgtcatcttcactctgtgcctacaactcccccacactcgaagcagtg	358
Db	304	CTTGGGGAAAAGAAAGTGTATTACTTGTGTGATTCMAAAATCCCAACATTGAGCACAAG	363
QY	359	acttcacagtctcggatgataaggaatgtcttcctgaagctgtctcgtcctcgcctactcg	418
Db	364	GCTATTCTTGAGCAACAAAGAGAAAGAAAGAAAGTTCCTGCTGGCAATCTGCCCAAG	423
QY	419	ctgttact 426	
Db	424	CTGTAGCT 431	

RESULT 7
AF195545
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
914 bp mRNA linear PLN 01-NOV-2000
Arabidopsis thaliana putative histone deacetylase (HD2a) mRNA,
complete cds.
AF195545
AF195545.1 GI:11066134
.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosid II; Brassicales; Brassicaceae; Arabidopsi
s (bases 1 to 914)
Wu, K., Tian, L., Malik, K., Brown, D. and Miki, B.
Functional analysis of HD2 histone deacetylase homologs in
Arabidopsis thaliana

DEFINITION	AV539421 Arabidopsis thaliana roots Columbia Arabidopsis thaliana cDNA clone R2131f01F 3', mRNA sequence.
ACCESSION	AV539421
VERSION	AV539421.1
KEYWORDS	GI:8701178
SOURCE	EST.
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 426)
AUTHORS	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL	DNA Res. 7, 175-180 (2000)
MEDLINE	20363093
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/

TITLE	A large scale analysis of cDNA in <i>Arabidopsis thaliana</i> : Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL	RNA Res. 7, 175-180 (2000)
MEDLINE	20363093
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp , URL: http://www.kazusa.or.jp/en/plant/ .
FEATURES	Location/Qualifiers
SOURCE	1. 426

```

/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R2131f01F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/lssue_type="roots"
/notes=Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
111 a 103 c 74 g 138 t
BASE COUNT

```

```

BASE COUNT      111 a 103 c 74 g 138 t
ORIGIN

Query Match      35.0%; Score 426; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 3.4e-50;
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	755	cagcttctactctcgaagacagaagaagaagaagagacacacccgacacccac	814
Db	426	CAGCTTCTACTCTCGAGACAGAGAAGAAGAAGGAGACACACCGCCACGCAC	367
Qy	815	acccagctaaagaaggtggaaagtctctctgaatgtaaccagagaccccaagtctgag	874
Db	366	ACCCAGCTAAGAAGGCTGGAAAGTCTCCTGTGAATGCTAACAGAGACCCCAAGTCTGGAG	307

QY	675	gtccatcatccgctggtgatacaaacacaaagaagccattcaactcaggcaacaacatttggtg	934
Db	306	gtcaatcatccgctggtgatacaaacacaaagaagccattcaactcaggcaacaacatttggtg	247
QY	935	gtcccaacaacaagggtttctcaacaggggcaagggaagggtagcttaaggagcgtggat	994
Db	246	gttccaaacaacaagggtttctcaacaggggcaagggaagggttagcttaaggagcgtggat	187
QY	995	caaggagaggttttggtgttttcagtagatgatgaacaacacttggaaagtgtggttttggga	1054

Db	186	CAAGGAGAGGTTTTGGCTTTTCGAGTAGATGATGAACAACACTTGGAGTGTGTTTTGGA	127
QY	1055	tttttactttattttagtataaacttggtatcgatgagctattttgagttatgcaa	1114
Db	126	TTTTTATCTTATTTTATTAGTAATACTCTTTACGGATGAGCTATTTTGACTATTTGCCAA	67
QY	1115	tftctactttctctatgtaattcagttatatagaatatttgcgtgaatagagaagaagactcg	1174

DB 66 TTCTACITTCCTATGTAATTCAGTATATGAATATTTCCTGAAATGAGAAAGAAGACTCG 7

QY 1175 aattgc 1180
|||||

Db 6 AATTGC 1

20363093
 MEDLINE
 COMMENT
 Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Kazusa 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1. .477
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="SQ233q03F"
 /clone_lib="Arabidopsis thaliana green siliques Columbia"
 /issue_type="green siliques"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 117 a 113 c 90 g 157 t
 ORIGIN
 Query Match 39.0%; Score 475.4; DH 9; Length 477;
 Best Local Similarity 99.8%; Pred. No. 4.4e-57;
 Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 705 tgtatccaaacacccgtctctgaaagagcaaacacacagcagcagcagccltclac 764
 Db 477 TGTATCCAAAACACCGTCTCTCGAAGAAGGCAAAACACGACGACGACCTTCTAC 418
 QY 765 tctcagaagcagaagagaagaagaaagagagacacacccgcacacacccagctaa 824
 Db 417 TCCTCAGAAGCCAGAGAGAGAAGAAAGAGGAGGACACACCGCCACACACACCGCTAA 358
 QY 825 gaaggtgaaagtctcctgtgaatgctaaccagagagcccaagtctggaagtcacatc 884
 Db 357 GAAGGGTGGAAAGTCTCCTGTGAATGCTTAACAGAGCCCAAGTCTGGAGGTCAATCATC 298
 QY 885 cgggtgttaacacacaaagaagccattcaactcaggcacaacaatttggtagttccaaca 944
 Db 297 CGGTGTGAACAACAACAAGAGCCATTCACTCAGGCAACAATTTGGTGTCCAACAA 238
 QY 945 caaggttctacaagggcagaaggggaagggtagagcttaaggcgtggtacaaggagagg 1004
 Db 237 CAAGGGTCTTAACAAGGGCAAGGGAAGGGTAGAGCTTAAGGACGTGGATCAAGGAGG 178
 QY 1005 ttttggttttcagtagatgatgaacaacattggaagtggtgttttgatttatctt 1064
 Db 177 TTTTGGGTTTTCCAGTAGATGATGAACAACATTTGGAAGTGTGTTTGGATTTTATCTT 118
 QY 1065 attttatagtataacttgattcgatgaagctatttgaattatgcaalltctacttt 1124
 Db 117 ATTTTATTAGTATTAAGTGTGTTATCGGATGAGCTATTTTGAATTTGCAATTTCTACTTT 58
 QY 1125 cctatgaatttcagtagatgaattttgctgaatgagaagagactcgaaatgca 1181
 Db 57 CCTATGAATTCAGTATATGAATATTGCTGAATGAGAAGAAGACTCGAATTGCA 1
 RESULT 8
 AW004431/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 thale cress.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 585)
 Chen, J.; Momiya, M.; Chan, E.; Mooney, M.; Carroon, B.; Gilliland, D.;
 Wang, X.; Hillman, J.; Guebler, K.; Kim, C.; Dowle, M.; Brzoska, P.

QY 1142 atgaatatttctgaaatgagaagaactgcaattgcaaac 1184
 |||||||
 Db 45 ATGAATATTTCGTAATGAGAAGAAGACTCGAATTGCAAAC 3
 |||||||

RESULT 3
 LOCUS AV543549/c 559 bp mRNA linear EST 07-SEP-2000
 DEFINITION AV543549 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 CDNA clone R2202el2F 3', mRNA sequence.
 ACCESSION AV543549
 VERSION AV543549.1 GI:8714963
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 559)
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 JOURNAL DNA Res. 7, 175-180 (2000)
 MEDLINE 20363093
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
 source
 1..559
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="R2202el2F"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /tissue_type="roots"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"
 BASE COUNT 129 a 132 c 103 g 195 t
 ORIGIN

Query Match 43.9%; Score 534.8; DB 9; Length 559;
 Best Local Similarity 98.6%; Pred. No. 2.3e-65;
 Matches 550; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 628 gaagattctgagatgaagaggaggag-ctcttaagaagcctgagccaatcaaca 686
 |||||||
 Db 558 GGAGGATTCTGAGTATTAGAAGAGGAGGAGACTCTTAAGAAGCCTAAGCCCAATCAACAA 499
 |||||||

QY 687 gaagaggccaatgaatctgtatccaaacacccctctctggaagaggcaaacaccgc 746
 |||||||
 Db 498 AAGAGGCCAATGAATGTATGATCAAAACACCCCTCTCTGCAAGAGAGGCAAAACCCAGC 439
 |||||||

QY 747 aqacacccagcttctctctcaagacacagagagagagagagagagacacaccgc 806
 |||||||
 Db 438 AGCAGACACAGCTTCTACTCTCAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
 |||||||

QY 807 cacacacacccagctgaagagggtggaaagtctctgtgaatgtaaccagagagcccaa 866
 |||||||
 Db 378 CACACACACCCAGCTAAGAGAGGTGGAAAGTCTCTGTGAATGCTACCAAGAGGCCCA 319
 |||||||

QY 867 gctcagaggtcaatcatcccggtggttaacacacacacacacacacacacacacacac 926
 |||||||
 Db 318 CTCTGGAGGTCAATCATCTCGGTGGTAAACAACAACAACAACAACAACAACAACAACA 259
 |||||||

QY 927 atttggtgttccaaacaaaggggtcttaacaaaggggcaaggggtagagcttaagg 986
 |||||||
 Db 258 ATTTGGTGGTCCACACACAGAGGTTCTTAACAGGGCAGGGAAGGAGTAGGCTTAGG 199
 |||||||

QY 987 acgtggatcaagggaggggttttgggttttcagtagatgatgaataaacacttgggaagtgtg 1046
 |||||||

Db 198 ACCTGGATCAAGGAGAGGTTTTCGGTTCAGTAGATGATGAACACATGGGAAGTCTG 139
 |||||||

QY 1047 gtttggaattttatctattttatagtaaaccttgttatcgagatgaagctatlltgaqt 1106
 |||||||

Db 138 CTTTGGGATTTTATCTTATTATTATTAACCTTGTTATCGGATGAGCTATTTTCAGT 79
 |||||||

QY 1107 atttgcatttctacttctctatgaattcagttatgaataatttctgaaatgagaag 1166
 |||||||

Db 78 ATTTGCAATTTCTACTTTCTCTATGTAATTCAGTATATGATATTTCCTGAATGAGAAA 19
 |||||||

QY 1167 aagactggaattgcaaac 1184
 |||||||

Db 18 AAGACTCGAATTGCAAC 1
 |||||||

RESULT 4
 A1998449/c 552 bp mRNA linear EST 08-SEP-1999
 LOCUS thaliana CDNA clone 701545727, mRNA sequence.
 DEFINITION 701545727 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis
 thaliana CDNA clone 701545727, mRNA sequence.
 ACCESSION A1998449
 VERSION A1998449.1 GI:5845354
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 552)
 AUTHORS Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
 Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Arzoska, P.,
 Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D., Tan, R.,
 Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, T.,
 Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L.,
 Turner, C., Krikorian, S., Elder, L. and Hanson, D.
 TITLE Arabidopsis thaliana Gene Expression MicroArray
 JOURNAL Unpublished (1999)
 COMMENT Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyle
 Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.
 Location/Qualifiers
 1..552
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701545727"
 /clone_lib="A. thaliana, Columbia Col-0, rosette-2"
 /dev_stage="4 - 7 weeks"
 /tissue_type="rosette"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
 library was derived from untreated rosette tissue from
 Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.
 Plants were grown in 1:1:1 peat moss/vermiculite/perlite
 soil at 22 deg. C +/- 3 deg. C under constant light, and
 watered with fertilizer. cDNA synthesis was initiated
 using a NotI-oligo(dT) primer. Double-stranded cDNA was
 blunt-ended, ligated to SalI adaptors, digested with NotI,
 size-selected, and cloned into the NotI and SalI sites of
 the pSPORT vector."
 the pSPORT vector."
 BASE COUNT 125 a 134 c 92 g 178 t 23 others
 ORIGIN

Query Match 41.4%; Score 504.4; DB 9; Length 552;
 Best Local Similarity 95.3%; Pred. No. 3.9e-61;
 Matches 527; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY 598 gatgaagatgactcagatgatgacgaggagggaaggttctgaggatgaagaggaggga 657
 |||||||

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 18:43:44 ; Search time 3265.32 Seconds
(without alignments)
5034.512 Million cell updates/sec

Title: us-09-645-337-7
Perfect score: 1218
Sequence: 1 gcttcgcttctaaaaaa.....aaaaaaagggcgccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estab:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpi:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	619.2	50.8	728	10 BE844690	BE844690 AD01A07T7
C 2	554.8	45.6	584	9 AV543093	AV543093 AV543093
C 3	534.8	43.9	559	9 AV543549	AV543549 AV543549
C 4	504.4	41.4	552	9 AI998449	AI998449 701345727
C 5	502	41.2	503	9 AV542413	AV542413 AV542413
C 6	479.8	39.4	483	9 AV559263	AV559263 AV559263
C 7	475.4	39.0	477	9 AV565954	AV565954 AV565954
C 8	458.4	37.6	585	9 AW004431	AW004431 701931423
C 9	426	35.0	426	9 AV539421	AV539421 AV539421
C 10	412.4	33.9	422	10 BE523968	BE523968 M44A9STM
C 11	411.4	33.8	421	9 AV542939	AV542939 AV542939
C 12	409.4	33.6	504	10 T45874	T45874 9137 Lambda
C 13	407	33.4	444	9 AA597939	AA597939 29188 lam
C 14	406.4	33.4	422	9 AV537871	AV537871 AV537871
C 15	404	33.2	407	9 AV541169	AV541169 AV541169
C 16	383	31.4	383	9 AV550412	AV550412 AV550412
C 17	376.4	30.9	378	9 AV551556	AV551556 AV551556

18	363.4	29.8	382	10 C99809	C99809 C99809 YAC
19	354.6	29.1	457	9 AA713170	AA713170 32730 lam
C 20	351	28.8	387	9 AV539161	AV539161 AV539161
C 21	327	26.8	327	9 AV549937	AV549937 AV549937
C 22	274.6	22.5	432	9 AA395217	AA395217 27000 lam
C 23	267.6	22.0	282	9 AV539637	AV539637 AV539637
C 24	263	21.6	265	10 Z35338	Z35338 ATTS3777 AC
C 25	249	20.4	249	10 BE525763	BE525763 M63018STM
C 26	238.8	19.6	298	10 Z34946	Z34946 ATTS3623 ST
C 27	226	18.6	237	10 Z35337	Z35337 ATTS3776 AC
C 28	221.2	18.2	333	10 Z24526	Z24526 ATTS968 Ral
C 29	213	17.5	263	10 Z17912	Z17912 ATTS0396 AC
C 30	194	15.9	288	10 T43469	T43469 6732 Lambda
C 31	184	15.1	220	10 Z17565	Z17565 ATTS0139 AC
C 32	147.4	12.1	534	9 AI995198	AI995198 701502609
C 33	147.4	12.1	715	10 BF270691	BF270691 GA_Eb000
C 34	142.2	11.7	543	10 BF279194	BF279194 GA_Eb003
C 35	139.2	11.4	434	10 N37851	N37851 19078 Lambda
C 36	139.2	11.4	875	10 BF275467	BF275467 GA_Eb002
C 37	135.4	11.1	865	10 BG445639	BG445639 GA_Ea002
C 38	125.2	10.3	390	9 AU068818	AU068818 AU068818
C 39	125.2	10.3	517	10 BF278973	BF278973 GA_Eb003
C 40	123.2	10.1	458	10 C72062	C72062 C72062 Rice
C 41	122.4	10.0	787	9 AU082425	AU082425 AU082425
C 42	121.2	10.0	397	9 AU067990	AU067990 AU067990
C 43	117.8	9.7	692	10 BI543491	BI543491 S2S3A_H3
C 44	114.8	9.4	742	10 BG526132	BG526132 57-3 Stev
C 45	113.8	9.3	402	10 D15380	D15380 RICC0557A R

ALIGNMENTS

RESULT 1

BE844690
LOCUS
DEFINITION
AD01A07T7 AD A. thaliana (Col-0 gll) library enriched for salt-induced transcripts; 10-14 day seedlings; 4h 160mM NaCl stress
Arabidopsis thaliana cDNA clone AD01A07 similar to (AF044914)
putative histone deacetylase, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
BE844690
BE844690.1 GI:10277068
EST.
thale cress.

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 728)
Gong, Z., Kolwa, H., Cushman, M.A., Ray, A., Bufford, D., Kore-Eda, S., Matsumoto, T.K., Zhu, J., Cushman, J.C., Bressan, R.A. and Hasegawa, P.M.

TITLE
Genes that are uniquely stress regulated in salt overly sensitive (sos) mutants

JOURNAL
MEDLINE
Plant Physiol. 126 (1), 363-375 (2001)
21249177

COMMENT

Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Plate: AD01 row: A Column: 7
Seq primer: T7
POLY-A-No.

FEATURES

Location/Qualifiers
source
1..728
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia gll"
/db_xref="taxon:3702"
/clone="AD01A07"
/clone_lib="AD A. thaliana (Col-0 gll) library enriched


```

Query Match          4.1%; Score 49.8; DB 3; Length 397;
Best Local Similarity 49.8%; Pred. No. 0.0016;
Matches 126; Conservative 0; Mismatches 127; Indels 0; Gaps
QY 365 etagtctgagatgagagatgtctctgaagctgttctctctctctctctctctactgtgta 420
   || || || || || || || || || || || || || || || || || || || || || ||
Db 322 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 263
QY 425 ctgccaaagaaatgctggagcagctgttctcaagctgacacaaagccaaagccaaac 480
   || || || || || || || || || || || || || || || || || || || || || ||
Db 262 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203
QY 485 ctgccgaagtgaagcctgcgaagagagaagcctgaatccagacagagaagatgagctcatg 540
   || || || || || || || || || || || || || || || || || || || || || ||
Db 202 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 143
QY 545 atgaagatgaagctgaagagagatgactcttgagaaggaatggatggttcatgaa 600
   || || || || || || || || || || || || || || || || || || || || || ||
Db 142 CTGCTGTTGCTGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83
QY 605 atgactcagatga 617
   || || || || || || || || || || || || || || || || || || || || || ||
Db 82 AAAACTTAAAGA 70

RESULT 12
US-08-738-349-1
; Sequence 1, Application US/08738349
; Patent No. 5869638
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Okazaki, Makoto
; APPLICANT: Kawai, Shinji
; APPLICANT: Tsujimura, Atsushi
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Cadherin-Like Protein and
; TITLE OF INVENTION: Process for Its Production
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,349
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/364,439
; FILING DATE:
; APPLICATION NUMBER: US 08/112,061
; FILING DATE: 26-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barker, M. P.
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 02481.1323-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3581 base pairs
; TYPE: nucleic acid

```

```

RESULT 9
US-08-766-738-2
; Sequence 2, Application US/08766738
; Patent No. 5916749
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHATASE INHIBITOR PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

Qy 674 agccaatcaacaagaagaggccaatga 701
 || - - - ||| ||| ||| |||
 Db 831 AGGAGAAGATGATTAAgACCCCAATGA 858

```

RESULT 10
US-07-867-106-2/c
: Sequence 2, Application US/07867106
: Patent No. 5389526
: GENERAL INFORMATION:
: APPLICANT: Slade, Martin B
: APPLICANT: Chang, Andy C M
: APPLICANT: Williams, Keith L
: TITLE OF INVENTION: Improved plasmid vectors for Cellular
: TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 53895261rs
: STREET: One Liberty Place 46th Floor
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103

```

```
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgt-F1s
US-08-232-463-14

Query Match 7.0%; Score 85.4; DB 1; Length 7218;
Best Local Similarity 6.2%; Pred. No. 6.8e-12;
Matches 26; Conservative 246; Mismatches 147; Indels 0; Gaps 0;

QY 422 ttactgcacagaaatgctggagcagctgttgcgaagctgacacaaagccaaagccca 481
Db 1441 TGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382
QY 482 aacctgccaaatgaagcctgcagagaaagcctgaatcacagcagaggaagatgactcg 541
Db 1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322
QY 542 atgatgaagatgaagctcgaagagatgatgactcttgagaagaatggatggtgatg 601
Db 1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262
QY 602 aagatgactcagatgatcagcagaggaggaagattctgagatgaagaagagagagctc 661
Db 1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202
QY 662 ttaagagcctgaagcacaagagagccaaatgaatctgtatccaaacacccg 721
Db 1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142
QY 722 tctctggaagaaggcaaacaccagcagcagaccagcttctactcagaagacagaag 781
Db 1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082
QY 782 agaagaagaagagagacacccgcacacacacccagctaaaggggtgaaagtct 840
Db 1081 RRRRRRRRRRRRRRRATCGAAGCTCCCTCGACCTGCAGCCAAAGCTCGAATTAATCT 1023

RESULT 6
US-09-177-325-2
; Sequence 2, Application US/09177325B
; Patent No. 6214983
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sakin, Ugur
; APPLICANT: Pfeundschuh, Michael
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/09/177,325B
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 2
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-177-325-2
```

```
Query Match 4.7%; Score 57.8; DB 4; Length 1276;
Best Local Similarity 51.3%; Pred. No. 2.5e-05;
Matches 134; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 373 gatgatgaggatgttctctgaagctgttctctgctctctctctctctctctctctct 432
Db 645 gatgaagcagaagaagcagacgcagcagcagcagcagcagcagcagcagcagcagaa 704
QY 433 ggaatgctggagcagctgttgcgaagcctgaatcacagcagaggaagatgactgaa 492
Db 705 gccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 764
QY 493 gtgaagcctgcagaaagagaaagcctgaatcacagcagaggaagatgactgaa 552
Db 765 gagggaggaggaagaagaagaggaaggaagaagaagaagaagaagaagaaga 824
QY 553 gaagtctgaagagatgatgactcttgagaagaagaatggatggttgatgaagatgactca 612
Db 825 ggagaagaagagagagagagagagagagagagagagagagagagagagagagaa 884
QY 613 gatgatcagcagaggaggaagga 633
Db 885 gaagaagaggaggaaggaaga 905

RESULT 7
US-09-411-812A-2
; Sequence 2, Application US/09411812A
; Patent No. 6261778
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sakin, Ugur
; APPLICANT: Pfeundschuh, Michael
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/09/411,812A
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: US 09/177,325
; PRIOR FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-411-812A-2

Query Match 4.7%; Score 57.8; DB 4; Length 1276;
Best Local Similarity 51.3%; Pred. No. 2.5e-05;
Matches 134; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 373 gatgatgaggatgttctctgaagctgttctctgctctctctctctctctctctctct 432
Db 645 gatgaagcagaagaagcagacgcagcagcagcagcagcagcagcagcagcagcagaa 704
QY 433 ggaatgctggagcagctgttgcgaagcctgaatcacagcagaggaagatgactgaa 492
Db 705 gccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 764
QY 493 gtgaagcctgcagaaagagaaagcctgaatcacagcagaggaagatgactgaa 552
Db 765 gagggaggaggaagaagaagaggaaggaagaagaagaagaagaagaagaagaaga 824
QY 553 gaagtctgaagagatgatgactcttgagaagaagaatggatggttgatgaagatgactca 612
Db 825 ggagaagaagagagagagagagagagagagagagagagagagagagagagagaa 884
QY 613 gatgatcagcagaggaggaagga 633
Db 885 gaagaagaggaggaaggaaga 905
```

QY 313 aatgcttattcattgctacaaalcccccaacatocagcagcagatgacttactatgct 372
 Db 333 agtattttttctgttctacaaagttgagcagcattgagggagatgaaatggatctt 392
 QY 373 gatgatgagatgttctgaagctgttctgctctg-----ccctactgct 420
 Db 393 gattctgagatgaagaggagagctaaacalccagtaatacaggaataatggcaagct 452
 QY 421 gttactgccacgaaatgctggagcagctgttgcaagctgttgcaagcgaacgaagccc 480
 Db 453 gatgggaaggaggagcagaataatcaagaagaaggcagctgcttactacagcttcaaatca 512
 QY 481 aaacctgccaaagtgagcctgcagaagaagaagcctgaatcaacacgaggaagatgact 540
 Db 513 agtctgcttgaagaagaagaagcagagctgactctgctgattctgagatgagctgct 572
 QY 541 gatgatgaagatgaagctgaagaggatgacttctgagaaaggaatggat----- 593
 Db 573 gatgattctatga---gagatattctgattctgattctgagaaaggcagggattatctcc 628
 QY 594 ---ggttgatgaagatgactcagatgacgagaggaggaaggttctgagatgaagaag 650
 Db 629 tgacgaaggcagatgattcaagtgatgaggtgataccagtgatgacgaggaaga 688
 QY 651 agggaggcttcttaagaagcctgagccatcaacaagaagaagcccaaatgaatctgtatc 710
 Db 689 aaccccaactcctaaagcagagga---ggcaagaagagaggtgctgaaatgctct 745
 QY 711 caaaacacccgtctctgaaagaaggaagcaaacacagcagcagccagcttctactctca 770
 Db 746 gaaacaccccttctctgataagaagaagcaaggttgccacaccgag-----ccca 796
 QY 771 gaagcagaagaagaagaagaaggaggaacacacccgacacacccacacccagctaa 824
 Db 797 gaaacagggtggcaagaagggtgccaccatgtggcaactccacacccagcaaa 850

RESULT 4
 US-09-282-305-11
 : Sequence 11, Application US/09282305
 : Patent No. 6287843
 : GENERAL INFORMATION:
 : APPLICANT: Baldwin, Donald A.
 : APPLICANT: Briggs, Steven P.
 : APPLICANT: Crane, Virginia C.
 : TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 : FILE REFERENCE: 5718-44
 : CURRENT APPLICATION NUMBER: US/09/282,305
 : CURRENT FILING DATE: 1999-03-31
 : PRIOR APPLICATION NUMBER: 60/080,563
 : PRIOR FILING DATE: 1998-04-03
 : NUMBER OF SEQ ID NOS: 18
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 11
 : LENGTH: 1283
 : TYPE: DNA
 : ORGANISM: zea mays
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (79)..(996)
 US-09-282-305-11

Query Match 9.1%; Score 111.2; DB 4; Length 1283;
 Best Local Similarity 50.7%; Pred. No. 14e-18;
 Matches 359; Conservative 0; Mismatches 328; Indels 21; Gaps 3;

QY 61 atgaggtctggggaggtgcgggagacacaaagcctactaaggtgactcctgaaga 120
 Db 79 atgaggtctggggaggtgcgggagcctggttccactgttaaggtgagcctggtat 138
 QY 121 gacagccttccacattctcaggtctcacttgactgacagctgaaatcgagaatct 180

Db 139 ggcttctgctgcacctllccccaggctgctcttg---gggaalccgaagaagaglgalaal 195
 QY 181 gtggttttgaagtgactgtgtgtggggtctaaactgtttatggaaacactttcacagaac 240
 Db 196 gccttgatgtatgcataaattgatgcagaaacttgcctatggaaacctctctgttgac 255
 QY 241 aagttccctcagattagcttggtttlltggatlaaaagagtttggatllccacacagc 300
 Db 256 aagaaccacacatacaattgatctgatttgcgataaaagtttgagctgtgcacaca 315
 QY 301 ggtaccagaagcaaatcttcatttcattgctgcatacaaatcccccaacatcagcaggtgac 360
 Db 316 tcaaaactlaccagcttttcttctactggtgcacaaagttgagcagcattcgaggagaal 375
 QY 361 ttcactagttcgatgatgagatgttctgagctgttctgctctctgctccctactgct 420
 Db 376 gaaatggatctgattctgaagatgaagcggaggctgaatgttc-----cagta 426
 QY 421 gttactgccaaagcaatgctggagcagctgttctcaaggctgcacacaaagccaaagccc 480
 Db 427 gtcaggaaaaatggcaagctgatgagaagaacacagaagtcacagaagaagcaggttct 486
 QY 481 aaacctgcgaagtgagcctgcagaagaagagcctgaatcagac-----gaggaa 531
 Db 487 qcacctlcaaaatcaagtcggtatcccaagaagagcaggaatgacacgattctgatgag 546
 QY 532 gatgagctgagatgagatgaagctgaagaggtctgaagaggtgatgactcttgagaaaggatgg 591
 Db 547 gacgagctgagatgattctgatggatgagcggagcttctgatgaggtttgtctct 606
 QY 592 atggttgatgaagatgactcagatgatgcagaggaaggaaggaatcttgaagagaaaga 651
 Db 607 gaagaaggcagtgatgattcaagtgatgaagatgataccagtgacgatgagggaggaac 666
 QY 652 gggaggcttcttaagaagcctgagccaatcaacaagaagaagggcccaatgaatctatcc 711
 Db 667 actccaaactcctaaagcctgagglaggaagaagaagagcagcagagtcgagagtcg 726
 QY 712 aaaaacccgctctctgaaagaagcgaacacccagcagcagccagct 759
 Db 727 aaaaactctctatctgataagaagaagcagggttgccacaccgtctct 774

RESULT 5
 US-08-232-463-14/c
 : Sequence 14, Application US/08232463
 : Patent No. 5670367
 : GENERAL INFORMATION:
 : APPLICANT: DORNER, F.
 : APPLICANT: SCHEIFLINGER, F.
 : APPLICANT: FALKNER, P. G.
 : TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 : NUMBER OF SEQUENCES: 52
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: 1800 Diagonal Road, Suite 500
 : CITY: Alexandria
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22313-0299
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/232,463
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/935,313
 : FILING DATE:
 : APPLICATION NUMBER: EP 91 114 300.6

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

AAX90843	
ID	AAX90843 standard; DNA; 1191 BP.
XX AC	AAX90843;
XX DT	13-JAN-2000 (first entry)
XX DE	Maize histone deacetylase-7 DNA.
XX KW	Maize histone deacetylase; family 2, ZmHD2; chromatin structure;
XX KW	RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;
XX KW	transcription; plant transformation; heterochromatin; disease resistance;
XX KW	chromatin assembly; gene activity; toxin screening; pathogenicity;
XX KW	disease response promoter; ds.
XX OS	Zea mays.
XX FH	Key Location/Qualifiers
XX CDS	63..971
XX FT	/tag= a
XX FT	/product= "Maize histone deacetylase"
XX PN	W09951731-A2.
XX PD	14-OCT-1999.
XX PF	02-APR-1999; 99WO-US07370.
XX PR	03-APR-1998; 98US-0080563.
XX PA	(PION-) PIONEER HI-BRED INT INC.
XX PI	Baldwin DA, Briggs SP, Crane VC;
XX DR	WPI; 1999-611038/52.
XX PS	P-PSDB; AAY28803.
XX PT	New deacetylase genes, used for producing transgenic plants which have
XX PS	increased disease resistance -
XX PS	Claim 1; Page 73-75; 87pp; English.
XX CC	The present sequence encodes a maize histone deacetylase . This DNA
XX CC	belongs to family 2, ZmHD2 and appears to affect chromatin structure at
XX CC	promoters of RNA polymerase I and thus regulate ribosomal RNA production.
XX CC	The nucleotide sequence can be used to transform plants and increase
XX CC	disease resistance by altering the levels of heterochromatin, altering
XX CC	chromatin assembly, and gene activity of the transformed plants.
XX CC	Additionally, compositions find use in screening for toxins that affect
XX CC	pathogenicity and in determining which disease response promoters are
XX CC	regulated by histone deacetylase.
XX SQ	Sequence 1191 BP; 365 A; 242 C; 311 G; 273 T; 0 other;
Query Match 9.2%; Score 111.6; DB 20; Length 1191;	
Best Local Similarity 51.4%; Pred. No. 1.2e-15;	
Matches 362; Conservative 0; Mismatches 329; Indels 13; Gaps 4;	
QY	61 atggagttctgagggaattcgcgtgacacaaacacgactacttaagtgaactcctgaaga 120
DB	
QY	63 atggagttctgagggttcgagtgcaagcctggttccactgttaagtgtgagcctgat 122
QY	121 gacagccttgtcacattctcaggtctcactgacctgacagtgacaatctggagaatc 180
DB	
QY	123 ggctttgtctgcaccctccccagcgctcttg---gggaatcagaagaagtgataat 179
QY	
QY	181 gtggttttgagtgactgtgtgggggttaaactgttatgtggaacatttcacaagac 240
DB	
QY	180 gccttgatgtatgcataaatcatgatcagaacttgccattgggaaccctctctgttac 239
QY	241 aagltcccctcagatgtcttgattggtttttgataaagatttgagctttcacacacg 300
DB	

Db	240	aagaaccacacatttcaatttgcataaagatttgagctgctgcacaca	299
Qy	301	ggtaccaaacgaatgttcatttcattggtcctacaaaatcccccaacatcgacagatgac	360
Db	300	tcaaaaactaccagcgtcttcttccactggtcacaaggttgacacgcoattcgaggagat	359
Qy	361	ttcactagttcggatgagagatgttctcgaagctgttccctgctcctgccctactq--	418
Db	360	gaatggatcttgcatttcgaagatgaagacgagggcgtgaatgttccagtatcaaggaa	419
Qy	419	-ctgttactgcacacggaatgtcggagcagctgtgttcaaggctgacacaaagccaag	477
Db	420	aatdgcgaagctgatgggaagaacacagaaaaagtcgaagaaaggcagttgctgcacctca	479
Qy	478	gccaaacctgcgaagtgaacctgcagaagaagacgtgaatcagacgaggaagatgag	537
Db	480	aaatcaagctccggattccaagaagacgaagatgacgatgattgtgatgagcgagact	539
Qy	538	tctgatgatgaagatga--agtctgaagagatgalgactcttgagaaagaatlgalg	595
Db	540	gatgattctgatgagatgacgacgacgattctgatgaggtttatcttctgaagaa---	596
Qy	596	ttgatgaagatgactcagatgatgacgagggaggaggttctgaggtgaggaagaggag	655
Db	597	--ggcgatgatgattcaaagtgatgataccagtgacgaltgaggaggagacacac	654
Qy	656	gagctttaagaagcctgagccaatcaacaagaagagagccaaatgaatctgtatccaaa	715
Db	655	caactcctaagaagcctgagtgaggaagaagagagctgtgaaagttccgtgctgtaaa	714
Qy	716	caccctctcttgaaagaagggcaaaaccagcagcagcaccagct	759
Db	715	ctcctctatcataagaagaacaaaggttgcacacccgtcatct	758
RESULT 12			
AAX90844			
ID	AAX90844 standard; DNA; 1245 BP.		
XX AC	AAX90844;		
XX DT	13-JAN-2000 (first entry)		
XX DE	Maize histone deacetylase-8 DNA.		
XX KW	Maize histone deacetylase; family 2, ZmHD2; chromatin structure;		
XX KW	RNA polymerase I; ribosomal RNA production; promoter regulator; promoter;		
XX KW	transcription; plant transformation; heterochromatin; disease resistance;		
XX KW	chromatin assembly; gene activity; toxin screening; pathogenicity;		
XX KW	disease response promoter; ds.		
XX OS	Zea mays.		
XX FH	Key Location/Qualifiers		
XX CDS	84..1019		
XX FT	/*tag= a		
XX FT	/product= "Maize histone deacetylase"		
XX PN	W09951731-A2.		
XX PD	14-OCT-1999.		
XX PF	02-APR-1999; 99WO-US07370.		
XX PR	03-APR-1998; 98US-0080563.		
XX PA	(PION-) PIONEER HI-BRED INT INC.		
XX PI	Baldwin DA, Briggs SP, Crane VC;		
XX DR	WPI; 1999-611038/52.		
XX PS	P-PSDB; AAY28804.		

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132866.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155119.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

Db 302 cttgggaaaggaagtglttacttltgttgatatacaaaactcccaacattgagccacaag 361

Qy 359 acttcactagttcggatgatgaggatgttctcgaagctgttctgctcctcctcctactg 418

Db 362 gctattctgaggaagaagaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 421

Qy 419 ctgttact 426

Db 422 ctgtagct 429

RESULT 8

AAF80352

ID AAF80352 standard; DNA; 939 BP.

XX

AC AAF80352;

XX

DT 29-JUN-2001 (first entry)

XX

DE Nucleotide sequence of a histone deacetylase designated AthDA2A.

XX

KW Histone deacetylase; AthDA2A; RPD3; gene expression; transgenic plant;

KW HDAl; ethylene-responsive phenotype; hypocotyl elongation; ds.

XX

OS Arabidopsis thaliana.

XX

PH Key location/Qualifiers

FT CDS 49..786

FT /*tag= a

FT /product= "histone deacetylase AthDA2A"

XX

PN CA2316036-Al.

XX

PD 27-FEB-2001.

XX

PF 24-AUG-2000; 2000CA-2316036.

XX

PR 27-AUG-1999; 99US-0383971.

XX

PA (M1AC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX

PI Miki B, Brown D, Tian L, Wu K;

XX

DR WPI; 2001-258457/27.

DR P-PSDB; AAB67813.

PT Methods for regulating gene expression in transgenic plants, e.g.

PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl

PT elongation), comprises introducing genes encoding histone deacetylase

PT

XX

PS Claim 9; Fig 2A; 91pp; English.

XX

CC The present sequence encodes an Arabidopsis thaliana histone deacetylase

CC designated AthDA2A. The protein is homologous to yeast RPD3 and HDAl.

CC The polynucleotide sequence is used in the method of the invention.

CC The specification describes a method for regulating gene expression in

CC transgenic plants. The method comprises modifying histones by introducing

CC chimeric nucleotide sequences which have regulatory elements in operative

CC association with a gene of interest or with a nucleotide sequence

CC encoding histone deacetylase. The method is useful for regulating the

CC developmental, physiological or biochemical pathway within a plant,

CC particularly for repressing ethylene-responsive phenotypes

CC (e.g. inhibition of hypocotyl elongation). The method is also useful

CC as a functional test for identifying a phenotype associated with

CC perturbing a gene. The histone deacetylase genes are useful for

CC altering the development of an organism.

XX

SQ Sequence 939 BP; 296 A; 175 C; 236 G; 232 T; 0 other;

Matches 252; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

Qy 37 tctctctctctctctcaacaacaatggagttctggggagttgggtgacacacacacacac 96

Db 25 tttctcaaccttattcttagccatggagttctggggaaattgaagttaaatcagggaag 84

Qy 97 gctactaaggtagctcctgaagaagacagccttgcacatcttctcagggttcactt--- 153

Db 85 ccagttacagtgaactcctgaagaaggaagcattcttccacgtttctcaggcctgcttga 144

Qy 154 gactgcacagtgaatactctggagaatctgtgttttggagtgactgtgtgtgggctaaa 213

Db 145 gaatgtaaaacaagaaggaggaggttctgcttcttcttcttcttcttcttcttcttct 204

Qy 214 ctgttatttggacacacttccacaagacaaagttccctcctcagattagcttcttcttctt 273

Db 205 ttgtttctgggaactctatcagactgagaacatcctcctcagctttctctgtgtgttctt 264

Qy 274 gataaagatttggagcttctcacaacagcgttaccacaaagaaatgttcttcttcttctt 333

Db 265 gacaaggagtttggagcttctcacaacttggggaaaaggaggtttactttgttggatc 324

Qy 334 aaatcccccaacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 393

Db 325 aaactcccaacattgagccacaaggtcttctgaggaagaggaaggaaggaaggaagaa 384

Qy 394 gctgttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 426

Db 385 gttcctgctgggaatgctgccaaggctgtgact 417

RESULT 9

AAC43043

ID AAC43043 standard; DNA; 774 BP.

XX

AC AAC43043;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37816.

XX

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

PN EPI033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

Query Match 12.8%; Score 156.2; DB 22; Length 939;

Best Local Similarity 64.1%; Pred. No. 9.3e-26;


```

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.3%; Score 162.6; DB 21; Length 1002;
Best Local Similarity 57.1%; Pred. No. 3.4e-27;
Matches 316; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 2 tcttcgtcttcaaaaaaacctcaaacctctctctctctctctctctctctctctcaacaacaa 61
Db 4 tttcagctgcctcaaacctcaaacctctctctctctctctctctctctctctctctctctca 63
QY 62 tggagttctggggagttgcgtgacacacacacacacacacacacacacacacacacacacac 121
Db 64 tggagttctggggagttgcgtgacacacacacacacacacacacacacacacacacacacac 123
QY 122 acagccttgcacacacacacacacacacacacacacacacacacacacacacacacacacac 178
Db 124 gcatctcttccacacacacacacacacacacacacacacacacacacacacacacacacacac 183
QY 179 ctgtggtttgagtgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 238
Db 184 ttgcctttacatglaaagtgaggacacacacacacacacacacacacacacacacacacac 243
QY 239 acaagttccctcagatgattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 298
Db 244 agaacatccctcagcttctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 303
QY 299 gcgttacaaagcaaatgttcttcttcttcttcttcttcttcttcttcttcttcttcttct 358
Db 304 ctggggaagaaagaggtgttcttcttcttcttcttcttcttcttcttcttcttcttcttct 363
QY 359 acttcaactagtcgagtgatgagtgatgttctgagtgatgttctgagtgatgttctgagtgat 418
Db 364 gctattctgaggaagagaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 423
QY 419 ctgttactgcaacggaatgtgtgagcagctgttctcaaggctgacacaaagcacaagcacaag 478
Db 424 ctgccaaggtctgagctaaacacacacacacacacacacacacacacacacacacacacacac 483
QY 479 ccaaacctgcacagtgagcctgcagaaagaaagcctgaatcagacagaggaagatgaat 538
Db 484 atgaagaggaatgagctgattctgacggaatggatgaagatgattctgagtgaggatt 543
QY 539 ctgatgatgaaga 551
Db 544 ctgaggaagaaga 556

RESULT 7
AAC45610
ID AAC45610 standard; DNA; 991 BP.
XX
AC AAC45610;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47120.
XX

```

```

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

```

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 08-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142059.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

OY 740 aaccagcagcagcaccagcttctactctcagaagac 776
|||||
Db 432 aaccagcagcagcaccagcttctactctcagaagac 468

RESULT 5
AAC36931
ID AAC36931 standard; DNA; 940 BP.
XX AAC36931;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15573.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 25-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2002, 20:26:19 ; Search time 665.94 Seconds
(without alignments)
3140.226 Million cell updates/sec

Title: US-09-645-337-7

Perfect score: 1218
Sequence: 1 gcttgcgtctctaaaaa.....aaaaaaagggcgccgc 1218

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

```
1: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1139.8	93.6	1212	22	AAF80353
2	1036.6	85.1	1147	21	AA41072
3	422.8	34.7	707	21	AAC52092
4	369.4	30.3	468	21	AAC37832
5	165.2	13.6	940	21	AAC36931
6	162.6	13.3	1002	21	AAC35217
7	159.2	13.1	991	21	AAC45610
8	156.2	12.8	939	22	AAF80352
9	149.8	12.3	774	21	AA43043

10	112	9.2	1307	20	AA90845	Maize histone deac
11	111.6	9.2	1191	20	AA90843	Maize histone deac
12	111.2	9.1	1245	20	AA90844	Maize histone deac
13	111.2	9.1	1283	20	AA90842	Maize histone deac
14	110.4	9.1	331	21	AAC52921	Arabidopsis thalia
15	90	7.4	741	21	AAC40930	Arabidopsis thalia
16	66.2	5.4	856	21	AAC49668	Arabidopsis thalia
17	66.2	5.4	879	21	AAC34537	Arabidopsis thalia
18	59.8	4.9	690	23	AAS90703	DNA encoding novel
19	59.8	4.9	693	23	AAS74240	DNA encoding novel
20	59.8	4.9	693	23	AAS90715	DNA encoding novel
21	58.4	4.8	282	23	AAS90681	DNA encoding novel
22	57.8	4.7	360	22	ABA71815	Human foetal liver
23	57.8	4.7	360	22	AAK20192	Human brain expro
24	57.8	4.7	360	22	AAK46257	Human bone marrow
25	57.8	4.7	360	22	AAI52158	Probe #20844 used
26	57.8	4.7	1276	21	AAI15086	Transcript SCP-3H
27	56.8	4.7	425	22	AAH33149	Human colon cancer
28	56.6	4.6	1686	16	AAQ87587	DNA encoding Leuco
29	55.8	4.6	450	22	ABA59302	Human foetal liver
30	55.8	4.6	450	22	AAK07521	Human brain expro
31	55.8	4.6	450	22	AAK3313	Human bone marrow
32	55.8	4.6	450	22	AAI39102	Probe #7788 used
33	55.4	4.5	932	21	AAC59530	Human secreted pro
34	54.6	4.5	612	22	AAH71471	Human cervical can
35	54	4.4	2814	23	AAS65681	DNA encoding novel
36	53.6	4.4	411	22	AAI88595	Human polynucleoti
37	53.4	4.4	337	22	AAI11578	Human breast cance
38	52.8	4.3	351	22	AAI14805	Human breast cance
39	52.6	4.3	966	20	AA60314	DNA encoding a hum
40	52.4	4.3	198	22	AAS07727	Cervical cancer pr
41	52.4	4.3	393	23	AAS90684	DNA encoding novel
42	52.2	4.3	318	21	AAC10202	Human secreted pro
43	52	4.3	2334	23	AAS90705	DNA encoding novel
44	51.8	4.3	447	23	AAS75731	DNA encoding novel
45	51.8	4.3	510	23	AAS69539	DNA encoding novel

ALIGNMENTS

RESULT 1

AAF80353
ID AAF80353 standard; DNA; 1212 BP.

XX

AC AAF80353;

XX

DT 29-JUN-2001 (first entry)

XX

DE Nucleotide sequence of a histone deacetylase designated AHD2B.

XX

KW Histone deacetylase; AHD2B; RPD3; gene expression; transgenic plant;

KW HDA1; ethylene responsive phenotype; hypocotyl elongation; ds.

XX

OS Arabidopsis thaliana.

XX

FH Key Location/Qualifiers

CDS 61..978

FT /tag- a

FT /product- *histone deacetylase AHD2B*

XX

PN CA2316036-Al.

XX

PD 27-FEB-2001.

XX

PF 24-AUG-2000; 2000CA-2316036.

XX

PR 27-AUG-1999; 99US-0383971.

XX

PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX

PI Mikl B, Brown D, Tlan L, Wu K;


```

Db 121 TAACCTGTTATCGATGACCTATTTTTCAGTATTTTGGCAATTTCTACTTTCTCTATGTAATTC 62
QY 1137 agtatgaatatatttctgtaaatgagaagaagactcgaatgcacaaacaaa 1189
|||||
Db 61 AGTATGAAATTTTCTGCTGAAATGAGAAGAAGACTCGAATTGCAACACAAA 9

RESULT 12
T45874
LOCUS 504 bp mRNA linear EST 09-JAN-1998
DEFINITION 9137 Lambda-PRL2 Arabidopsis thaliana cDNA clone 133F15T7, mRNA
ACCESSION T45874
VERSION T45874.1 GI:934127
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 504)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
On Apr 14, 1993 this sequence version replaced gi:638462.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcm@ibm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 504
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA."
BASE COUNT 131 a 113 c 109 g 142 t 9 others
ORIGIN
|||||

Query Match 33.6% Score 409.4; DB 10; Length 504;
Best Local Similarity 95.9%; Pred. No. 6.4e-48;
Matches 439; Conservative 0; Mismatches 16; Indels 3; Gaps 2;

QY 2 tcttcgctcttaaaaaaaccttaaacctctctctctctctctctcgttcacacaaa 61
|||||
Db 1 TCTTTCCTCTATAAAAAACCTTANCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY 62 tggagttctgggggagttcgcggtgacacacacacacacacacacacacacacacac 121
|||||
Db 61 TGGAGTTCTGGGGAGTTGCGGTGACACACACACACACACACACACACACACACAC 120
QY 122 acagcctgtccacattctcaggttcacttcactgcacaggaatctggagaatctg 181
|||||

```

```

Db 121 ACAGCCTGTCCACATTTCTCAGGCTTCTACTTGACATGACAGTGAATCTGGAGAAATCTG 180
QY 182 tggtttgagtgactgttgggtggggcctaaacttggatggacacactlllcacaaagaca 241
|||||
Db 181 TGGTTTGGAGTGTGACTGTTGGTGGGGCTAAACTTGTATTGGAACACTTTTACAAGACA 240
QY 242 agttccctcagattagcttggattggttttttgataaagagtttgagctttcacacacagcg 301
|||||
Db 241 AGTTCCCTCAGATTAGCTTTGATTGTTTGTATTAAGAGTTTGAGCTTTTACACACAGCG 300
QY 302 gtacaaaagcaaatgcttctcattcattgctgacaaalcccccaacatcgcagcaggaagac 361
|||||
Db 301 GTACAAAGCAAAATTTTCATTTCTTGGCTTACAAATCCCCCAACATCGGACGAGGATGACT 360
QY 362 tcaactagctcggatgagagatgttccctgaagcgttctcctcctcctcctcctcctcctc 421
|||||
Db 361 TCACTAGTTCGGATGATGAGGATGTTCTCTGGAAGCTTCTCTGTCTGNCCTCC--ACTGNTG 418
QY 422 ttactgccaacgaaatgctggagcagctgtgtgtcgaag 459
|||||
Db 419 TTACTGCC-ACGGGAATCCTGGGGCAGCTTTTTCNCAAGG 455

RESULT 13
AA597939 444 bp mRNA linear EST 19-SEP-1997
LOCUS 29188 Lambda-PRL2 Arabidopsis thaliana cDNA clone 248C11T7, mRNA
DEFINITION
ACCESSION AA597939
VERSION AA597939.1 GI:2413362
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 444)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcm@ibm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers
1. 444
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dT primed cDNA."
BASE COUNT 138 a 84 c 109 g 106 t 7 others
ORIGIN
|||||

```


0000000000

BASE COUNT

BASE COUNT

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:03:52 ; Search time 144.83 Seconds
(without alignments)
233.912 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWCVATPNATKVTPEE.....KQFGSNKNGKKGKGRA 305

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /net/abs06/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	305	22	AA667814
2	1558	98.0	305	21	AAG26185
3	740	46.6	155	21	AAG17704
4	594	37.4	245	21	AAG38221
5	594	37.4	245	22	AA667813
6	594	37.4	257	21	AA631484
7	589.5	37.1	248	21	AAG10894
8	544	34.2	311	20	AA228804
9	535	33.7	302	20	AA228803
10	528	33.2	305	20	AA228802
11	436	27.4	285	20	AA228805

12	393.5	24.8	149	21	AAG54981	Arabidopsis thalia
13	345.5	21.7	208	21	AAG25823	Arabidopsis thalia
14	311	19.6	412	20	AAW68010	S. frugiperda immu
15	262.5	16.5	203	21	AAG09171	Arabidopsis thalia
16	227	14.3	349	22	AB662190	Drosophila melanog
17	223.5	14.1	194	21	AAG54925	Arabidopsis thalia
18	219	13.8	98	21	AAG15354	Arabidopsis thalia
19	216.5	13.6	181	21	AAG09172	Arabidopsis thalia
20	212.5	13.4	126	21	AAG15353	Arabidopsis thalia
21	205	12.9	707	16	AAW79912	Human nucleolin.
22	205	12.9	707	16	AAW84052	Human V3 loop HIV
23	205	12.9	707	22	AAW84052	Human nucleolin.
24	200.5	12.6	570	22	AB657851	Drosophila melanog
25	194.5	12.2	654	22	AB663266	Drosophila melanog
26	194.5	12.2	688	22	AB663269	Drosophila melanog
27	193.5	12.2	147	21	AAG09173	Arabidopsis thalia
28	190.5	12.0	536	22	AAW92836	Human protein sequ
29	189	11.9	857	22	AB663922	Drosophila melanog
30	188	11.8	723	14	AAW38745	Drosophila SSRP co
31	188	11.8	723	19	AAW39213	D. melanogaster SS
32	188	11.8	723	22	AB661000	Drosophila melanog
33	184.5	11.6	682	21	AAG46506	Arabidopsis thalia
34	184.5	11.6	724	21	AAG46505	Arabidopsis thalia
35	184.5	11.6	778	21	AAG46504	Arabidopsis thalia
36	180	11.3	1162	21	AAW96255	Kaposi's sarcoma-a
37	180	11.3	1162	21	AAW58500	HIV8 ORF 73 protei
38	180	11.3	1162	22	AAW62331	Amino acid sequenc
39	179.5	11.3	670	21	AAG51795	Arabidopsis thalia
40	179.5	11.3	688	21	AAG51794	Arabidopsis thalia
41	179.5	11.3	722	21	AAG51793	Arabidopsis thalia
42	178.5	11.2	517	22	AAW39867	Human polypeptide
43	178.5	11.2	520	22	AAW41653	Human polypeptide
44	177.5	11.2	885	21	AAW18279	Plasmodium falcipa
45	174.5	11.0	718	21	AAW42569	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB67814
ID AAB67814 standard; Protein; 305 AA.
XX AAB67814;
AC AAB67814;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated ALHDA2B.
XX
DE Histone deacetylase; ALHDA2B; RPD3; gene expression; transgenic plant;
KW HDAL; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
PN CA2316036-Al.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX

(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

PI Miki B, Brown D, Tian L, Wu K;

DR WPI; 2001-258457/27.

DR N-PSDB; AAF80353.

PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase

XX Claim 10; Fig 2B; 91pp; English.
 CC The present sequence represents Arabidopsis thaliana histone deacetylase
 CC designated AT4G22820. The protein is homologous to yeast RPD3 and HDAC1.
 CC The polynucleotide sequence is used in the method of the invention.
 CC The specification describes a method for regulating gene expression in
 CC transgenic plants. The method comprises modifying histones by introducing
 CC chimeric nucleotide sequences which have regulatory elements in operative
 CC association with a gene of interest or with a nucleotide sequence
 CC encoding histone deacetylase. The method is useful for regulating the
 CC developmental, physiological or biochemical pathway within a plant,
 CC particularly for repressing ethylene-responsive phenotypes
 CC (e.g. inhibition of hypocotyl elongation). The method is also useful
 CC as a functional test for identifying a phenotype associated with
 CC perturbing a gene. The histone deacetylase genes are useful for
 CC altering the development of an organism.

XX Sequence 305 AA;

Query Match 100.0%; Score 1589; DB 22; Length 305;
 Best Local Similarity 100.0%; Pred. No. 4.7e-119;
 Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFMGVAVTPKNAKVTPEEDSLVHISQASLDCTVKSGESVLSVTGGAKLVIGTISQD 60
 Db 1 mefmgvavtpknatkvtpeedsllvhisqasldctvksesvlsvtvggkvlvigtlsqd 60
 QY 61 KFPQISDLVDFKFEFLSHSGTKANVHIFIGYKSPNIEQDDFTSSDDEVDPEAVPAPATA 120
 Db 61 kfpqisdldvdfkfehlshsgtkanvfhfigykspnieqddftssddedvdpavpapata 120
 QY 121 VTANGAGAAVVKADTKPKAPAEVKAPEEKPEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 180
 Db 121 vtangagaaavvkadtppkapaevkapeekpdeesddesddesddesddesddesddesddes 180
 QY 181 DSDDE 240
 Db 181 ddsddeeedsedeetpkpkpepknkpnsvsktpvsgkkaakpaaapastpqktek 240
 QY 241 KKGHTATPHPAKGGKSPVNAVSPKSGGSGGNKPKPFNSGKOPFGSGNNKSGNKGK 300
 Db 241 kkghtatphpakkggkspvnavspksgsggnnknkpfnsqkqfgsgnnksgnkgk 300
 QY 301 GKGRA 305
 Db 301 gkgra 305

RESULT 2

AAG26185
 ID AAG26185 standard; Protein; 305 AA.

AC AAG26185;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 30547.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.

KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.	PR	23-JUN-1999;	990S-0140354
KW		PR	24-JUN-1999;	990S-0140695
KW		PR	28-JUN-1999;	990S-0140823
XX		PR	29-JUN-1999;	990S-0140991
OS		PR	30-JUN-1999;	990S-0141287
XX	Arabidopsis thaliana.	PR	01-JUL-1999;	990S-0141842
XX	EP1033405-A2.	PR	01-JUL-1999;	990S-0142154
XX		PR	02-JUL-1999;	990S-0142055
PD	06-SEP-2000.	PR	06-JUL-1999;	990S-0142390
XX		PR	08-JUL-1999;	990S-0142803
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	990S-0142920
XX		PR	12-JUL-1999;	990S-0142977
XX	25-FEB-1999;	PR	13-JUL-1999;	990S-0143542
PR	05-MAR-1999;	PR	14-JUL-1999;	990S-0143624
PR	09-MAR-1999;	PR	15-JUL-1999;	990S-0144005
PR	23-MAR-1999;	PR	15-JUL-1999;	990S-0144085
PR	25-MAR-1999;	PR	16-JUL-1999;	990S-0144085
PR	29-MAR-1999;	PR	16-JUL-1999;	990S-0144086
PR	01-APR-1999;	PR	19-JUL-1999;	990S-0144325
PR	06-APR-1999;	PR	19-JUL-1999;	990S-0144333
PR	08-APR-1999;	PR	19-JUL-1999;	990S-0144332
PR	16-APR-1999;	PR	19-JUL-1999;	990S-0144333
PR	19-APR-1999;	PR	19-JUL-1999;	990S-0144334
PR	21-APR-1999;	PR	19-JUL-1999;	990S-0144335
PR	23-APR-1999;	PR	20-JUL-1999;	990S-0144352
PR	23-APR-1999;	PR	20-JUL-1999;	990S-0144632
PR	28-APR-1999;	PR	20-JUL-1999;	990S-0144884
PR	30-APR-1999;	PR	21-JUL-1999;	990S-0144814
PR	30-APR-1999;	PR	21-JUL-1999;	990S-0145086
PR	04-MAY-1999;	PR	21-JUL-1999;	990S-0145088
PR	05-MAY-1999;	PR	22-JUL-1999;	990S-0145085
PR	06-MAY-1999;	PR	22-JUL-1999;	990S-0145087
PR	06-MAY-1999;	PR	22-JUL-1999;	990S-0145089
PR	07-MAY-1999;	PR	22-JUL-1999;	990S-0145192
PR	11-MAY-1999;	PR	23-JUL-1999;	990S-0145145
PR	14-MAY-1999;	PR	23-JUL-1999;	990S-0145218
PR	14-MAY-1999;	PR	23-JUL-1999;	990S-0145224
PR	14-MAY-1999;	PR	26-JUL-1999;	990S-0145276
PR	14-MAY-1999;	PR	27-JUL-1999;	990S-0145913
PR	18-MAY-1999;	PR	27-JUL-1999;	990S-0145918
PR	18-MAY-1999;	PR	27-JUL-1999;	990S-0145919
PR	20-MAY-1999;	PR	28-JUL-1999;	990S-0145951
PR	21-MAY-1999;	PR	02-AUG-1999;	990S-0146386
PR	24-MAY-1999;	PR	02-AUG-1999;	990S-0146388
PR	25-MAY-1999;	PR	02-AUG-1999;	990S-0146389
PR	27-MAY-1999;	PR	03-AUG-1999;	990S-0147038
PR	28-MAY-1999;	PR	04-AUG-1999;	990S-0147204
PR	01-JUN-1999;	PR	04-AUG-1999;	990S-0147302
PR	03-JUN-1999;	PR	05-AUG-1999;	990S-0147192
PR	04-JUN-1999;	PR	05-AUG-1999;	990S-0147260
PR	07-JUN-1999;	PR	06-AUG-1999;	990S-0147303
PR	08-JUN-1999;	PR	06-AUG-1999;	990S-0147416
PR	10-JUN-1999;	PR	09-AUG-1999;	990S-0147493
PR	10-JUN-1999;	PR	09-AUG-1999;	990S-0147935
PR	14-JUN-1999;	PR	10-AUG-1999;	990S-0148171
PR	14-JUN-1999;	PR	11-AUG-1999;	990S-0148319
PR	16-JUN-1999;	PR	12-AUG-1999;	990S-0148341
PR	16-JUN-1999;	PR	13-AUG-1999;	990S-0148565
PR	17-JUN-1999;	PR	13-AUG-1999;	990S-0148684
PR	18-JUN-1999;	PR	16-AUG-1999;	990S-0149368
PR	18-JUN-1999;	PR	17-AUG-1999;	990S-0149175
PR	18-JUN-1999;	PR	18-AUG-1999;	990S-0149426
PR	18-JUN-1999;	PR	20-AUG-1999;	990S-0149722
PR	18-JUN-1999;	PR	20-AUG-1999;	990S-0149723
PR	18-JUN-1999;	PR	20-AUG-1999;	990S-0149929
PR	18-JUN-1999;	PR	23-AUG-1999;	990S-0149902
PR	18-JUN-1999;	PR	23-AUG-1999;	990S-0149930
PR	18-JUN-1999;	PR	25-AUG-	

PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143342.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 37.4%; Score 594; DB 21; Length 245;

Best Local Similarity 50.7%; Pred. No. 1e-39;

Matches 151; Conservative 26; Mismatches 63; Indels 59; Gaps 15;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHLSQASL-DCTVKSGSVLSTVTVGAKLVICTLSQ 59
 Db 1 mefwgievsgkgtvtlpeegllihvsqaslgcknkkgfvlphkvqanqlvqlat 60
 QY 60 DKPQISFDLVDFKPELSHSGTKANVHFIGYKSPNIEQDDFTSSDDVDVPEAVPAPAT 119
 Db 61 enipqlfcdlvdfkfelshsgtkgsvyfgkyttniepggy-sececeeevpa---- 115
 QY 120 AVTANGAGAAVVKADTKPKAKPAEVPKPESEDESEDESEDESEDESEDESEDESEKGMVD 179
 Db 116 -----gnaakav-----akpkapaevkpavd-----dcedesdsd-----gm--d 149
 QY 180 EDDSDDEDEDEDEDEDEDEETPKKPEPINKKRPNESYKTPVSCCKKAKPAAAPASTPQTF 239
 Db 150 eddsdg---eds---eceptpkpas-skkranetlpkapvvaakkakvav----tpqkld 199
 QY 240 KKKGGTATPHPAKKGCKSPVNAQSPKSGGSSGNNKPKFNSCKQPCGNNKGN 297
 Db 200 ek-----kkggka---angspksasqvscg-sckktfnsgnal-esnkhakh 241

[illegible]


```

Db      61  enipqlfcdlvdfkelfelstlwgkgyvfgyktpnlpepgy seeeeeeevpa ---- 115
Qy      120 AVTANGNAGAAVVKADTKPKAKPAEVKPAPEKPESEDESEDDSEKGMVD 179
Db      116  -----gnaakav-----akpkapaevkpvad-----deedesdad-----gm--d 149
Qy      180 EDDSDDEEEDSEDEEEETPKRPEPKRPNESVSKTPVSGKKAKPAAPASTPKOTE 239
Db      150 eddsdg---eds---eeceptkpkpas-skranettkpvksakkavav-----tpqktd 199
Qy      240 KKGCGHTATPHAKKGGKSPVNAQSPKSGQSGGNNKKPPNSGKQFGGNNKGSN 297
Db      200 ek-----kkgda-----anqspksasqvscg-sckktfnsnal-esmknakh 241

RESULT 7
AAG10894
ID  AAG10894 standard; Protein; 248 AA.
XX
AC  AAG10894;
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 9401.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD
XX
PE  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  28-APR-1999; 99US-0130891.
PR  30-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  04-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  07-MAY-1999; 99US-0132487.
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136392.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  03-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  10-JUN-1999; 99US-0138847.
PR  14-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  16-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139458.
PR  18-JUN-1999; 99US-0139459.
PR  18-JUN-1999; 99US-0139460.
PR  18-JUN-1999; 99US-0139461.
PR  18-JUN-1999; 99US-0139462.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  18-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  02-JUL-1999; 99US-0142154.
PR  06-JUL-1999; 99US-0142055.
PR  08-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  20-JUL-1999; 99US-0144352.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145087.
PR  22-JUL-1999; 99US-0145089.
PR  22-JUL-1999; 99US-0145192.
PR  23-JUL-1999; 99US-0145145.
PR  23-JUL-1999; 99US-0145218.
PR  23-JUL-1999; 99US-0145224.
PR  26-JUL-1999; 99US-0145276.
PR  27-JUL-1999; 99US-0145913.
PR  27-JUL-1999; 99US-0145918.
PR  27-JUL-1999; 99US-0145919.
PR  28-JUL-1999; 99US-0145951.
PR  02-AUG-1999; 99US-0146386.
PR  02-AUG-1999; 99US-0146388.
PR  02-AUG-1999; 99US-0146389.
PR  03-AUG-1999; 99US-0147038.
PR  04-AUG-1999; 99US-0147204.
PR  04-AUG-1999; 99US-0147302.
```

```

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 37.1%; Score 589.5; DB 21; Length 248;
Best Local Similarity 49.7%; Pred. No. 2.3e-39;
Matches 148; Conservative 27; Mismatches 68; Indels 55; Gaps 14;

QY 1 MEFWGVAATPKNATKTPPEEDSLVHISOASI-DCTVKSGESVVLVSVTVGGAKLVIGTISO 59
Db 1 mefwglevksgkpvvlveegililhsqslgecknkkgfvlphkvgnqnlvlgltst 60
QY 60 DKFPQISFDLVFDFKEFELSHSGTKANVHFITYKSPNIEQDDFTSSDDENVPPEAVPAPPT 119
Db 61 enlpqifcdlvdfkelfshsgtkgsvfygktpnlpoggysecececececepa-- 118
QY 120 AVTANGNAGAAVYKADTKPKAKPAEYKPAEKEPSEDESEDESEDESEDESEDESEKGMVD 179
Db 119 -----gnaakav-----akpkakpaevkavd-----deedesd-----gm--d 152
QY 180 EDDSDDEEDSDEREREFTPKKPEPINKKPNESVSKTPVSGKKAKKAPAAAPASTPQKTE 239
Db 153 eddsdg---eds---eeeeeipkpkpas--skkranetlpkpvssakkakvav---tpqkl 202
QY 240 KKKGGHTATPHPAKGGKSPVNAQSPKGGQSGGNNKPKPNNSCKQFGSGNNKGSN 297
Db 203 ek-----kkggka---anxspkasqvscg--skkttfsgnal--eshnkakh 244

RESULT 8
AAV28804
ID AAV28804 standard; protein; 311 AA.
XX
AC AAV28804;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-8.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
XX
OS Zea mays.
XX
PN MO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99MO-0507370.
XX
PR 03-APR-1998; 98US-0080563.
XX
(PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
WPI; 1999-611038/52.
DR N-PSDB; AAX90844.
XX
New deacetylase genes, used for producing transgenic plants which have
increased disease resistance
XX
Claim 1; Page 79-80; 87pp; English.
XX
The present sequence is maize histone deacetylase encoded by HD cDNA
belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
modifications, may be localised to promoters targeted for repression by
other proteins that associate with HD and specifically bind regulatory
elements in promoter DNA. The HD nucleotide sequence can be used for
producing transgenic plants with increased disease resistance.
CC
Additionally, compositions find use in screening for toxins that affect
pathogenicity and in determining which disease response promoters are
regulated by histone deacetylase.
XX
Sequence 311 AA;
SQ

```


CC other proteins that associate with HD and specifically bind regulatory
 CC elements in promoter DNA. The HD nucleotide sequence can be used for
 CC producing transgenic plants with increased disease resistance.
 CC Additionally, compositions find use in screening for toxins that affect
 CC pathogenicity and in determining which disease response promoters are
 CC regulated by histone deacetylase.

XX Sequence 305 AA;

Query Match 33.28; Score 528; DB 20; Length 305;
 Best Local Similarity 43.68; Pred. No. 2.4e-34;
 Matches 123; Conservative 44; Mismatches 93; Indels 22; Gaps 9;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISQASLDCTVKGESVLSVTVGGAKLVIGTISQD 60
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1 mefwglevkgstvkcepgygvfvlhlsqaalgeskks-dnalmyvkiddkilaigtlsvd 59

QY 61 KFOISFDLVDFKPEFELSHSGTKANVHFICYKSPNIFQDDFTSSDDSDVPEAVPAPATA 120
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 60 knphlgldlfdkelfshstktstsvftgkykveqfeedemldsdededeeinvp---v 116

QY 121 VTANCAAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESEED--DDSEKGMV 178
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 117 vkengkadekkkqekavapskspdksskdddsdedetdsdedetdsdedegls 176

QY 179 DEDSDDEDEDEDEDEDEDEE--TPKKPEPINKRPNESVSKTPVSGCKAKKAPAAAPASTPQ 236
 :
 Db 177 eegddsdeddtdaddeettpkkpevgkkraaesviktptlsdkkak-vatpss--q 233

QY 237 KTEKKKGK--HTATPHPAKKGKSPVNAOS-----PKSG 269
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 234 ktgkgkgaavhvavtphak--gktivnndksvkspkspksg 273

RESULT 11

AAY28805
 ID AAY28805 standard; protein: 285 AA.
 XX
 AC AAY28805;
 DT 13-JAN-2000 (first entry)
 XX
 DE Maize histone deacetylase-9.
 XX
 KW Maize histone deacetylase; HD; HD cDNA; family 2, ZmHD2; gene repression;
 KW acetyl modification; promoter; regulatory element; transgenic plant;
 KW disease resistance; toxin screening; pathogenicity;
 KW disease response promoter.
 XX
 OS Zea mays.
 XX
 PN WO951731-A2.
 XX
 PD 14-OCT-1999.
 XX
 PF 02-APR-1999; 99WO-US07370.
 XX
 PR 03-APR-1998; 98US-0080563.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Baldwin DA, Briggs SP, Crane VC;
 XX
 DR WPI; 1999-611038/52.
 DR N-PSDB; AAX90845.
 XX
 PT New deacetylase genes, used for producing transgenic plants which have
 PT increased disease resistance -
 XX
 PS Claim 1; Page 82-84; 87pp; English.
 XX
 CC The present sequence is maize histone deacetylase encoded by HD cDNA

CC belonging to family 2, ZmHD2. This enzyme responsible for removing acetyl
 CC modifications, may be localised to promoters targeted for repression by
 CC other proteins that associate with HD and specifically bind regulatory
 CC elements in promoter DNA. The HD nucleotide sequence can be used for
 CC producing transgenic plants with increased disease resistance.
 CC Additionally, compositions find use in screening for toxins that affect
 CC pathogenicity and in determining which disease response promoters are
 CC regulated by histone deacetylase.

XX Sequence 285 AA;

Query Match 27.48; Score 436; DB 20; Length 285;
 Best Local Similarity 36.08; Pred. No. 5e-27;
 Matches 114; Conservative 49; Mismatches 110; Indels 44; Gaps 11;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISQASLDCTVKGESVLSVTVGGAKLVIGTISQD 60
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 1 mefwglevkgstvkcepgygvfvlhlsqaalgeskks-dnalmyvkiddkilaigtlsvd 60

QY 61 KFOISFDLVDFKPEFELSHSGTKANVHFICYKSPNIFQDDFTSSDDSDVPEAVPAPATA 117
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 61 khpqiscdlvdfkelfshstktstsvftgkykveqfeedemldsdededeeinvp---v 117

QY 118 PTAVTANGAAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESEED--DDSEKGMV 177
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 118 ----mqnneikistakvkvvgiglnad---edetssgdddtfdsdseeds---d 167

QY 178 VDEDDSDDEDEDEDEDEDEE--TPKKPEPINKRPNESVSKTPVSGCKAKKAPAAAPASTPQ 236
 :
 Db 168 edevssdtldtsdsgseegtpkkttevvvgkkrale--aetp-sgkakk-----se 216

QY 237 KTEKKKGK--HTATPHPAKKGKSPVNAOS-----OSPKSGGSGSGNNKPPNNSGKOP 288
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 217 gsaqktgdkvstshpakgsktpadkstktpadkksksgshack-----scskaf 268

QY 289 GGSNNKSGSNKKGKGRA 305
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 Db 269 gsasaleshqkakkhea 285

RESULT 12

AAG54981
 ID AAG54981 standard; protein: 149 AA.
 XX
 AC AAG54981;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 70363.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.

PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	23-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	03-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	23-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150566
PR	26-AUG-1999	9905-0150884
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155139
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0158293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329

[illegible]

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.7%; Score 345.5; DB 21; Length 208;

Best Local Similarity 39.6%; Pred. No. 5.8e-20;

Matches 93; Conservative 27; Mismatches 66; Indels 49; Gaps 9;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISQASLCTVKSGES-----VVLSTVVGAKLV 53
||||| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 1 mefwgvavtkngkplhldpgldrlvlhlsqval-----gesknvntepqlgylvtvgadkl 54
QY 54 IGTLSODKFPQISFDLVFDEKPELSHSGTKANVHFITYKSPNIEODFTSSDDEDVPEAV 113
||||| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 55 igtlshekfplqstelveinfahtwngsvffggykv-----dasdpepdl---- 104
QY 114 PAPAPTAVTANGAGAAVYKADTKPKAK-----PAEVKPAEKPESD--EEDSDEDE 165
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 105 -----iddqleaagfkaapksaakqvnfqlpndvkaqddadgseedsddds 155
QY 166 SEEDDSEKGM---DVEDDDSDDEEEDSEDEEETPKKPEPINKK--RPNES 214
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 156 ensgdeeeekvtaesdeedsddds---seetpkpkeepkkrsaepnss 207

RESULT 14

AAW68010
ID AAW68010 standard; Protein; 412 AA.

XX AAW68010;

XX AC

XX DT 27-APR-1999 (first entry)

XX DE

XX S. frugiperda immunophilin FKBP46.

XX KW

XX KW Immunophilin; moth; insect cell; nuclear; immunosuppression; drug; transplant; tissue graft.

XX OS Spodoptera frugiperda.

XX FH Key Location/Qualifiers

FT Domain 111..114

FT FT /note- "EAP motif"

FT FT 124

FT FT /note- "putative nuclear localisation signal"

FT FT 135..138

FT FT /note- "putative nuclear localisation signal"

FT FT 148..149

FT FT /note- "putative nuclear localisation signal"

FT FT /note- "AP motif"

FT Domain 215..218

```

FT Region /note- "EEAP motif"
FT 219..223
FT /note- "putative nuclear localisation signal"
FT 250..256
FT /note- "putative nuclear localisation signal"
FT 272..275
FT /note- "putative nuclear localisation signal"
FT 284..286
FT /note- "putative nuclear localisation signal"
FT 292..295
FT /note- "EEAP motif"
FT 298..302
FT /note- "putative nuclear localisation signal"
XX PN US5861498-A.
XX 19-JAN-1999.
XX 31-OCT-1996; 96US-0741134.
XX 01-NOV-1995; 95US-0007163.
XX 31-OCT-1996; 96US-0741134.
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX Alnemri ES, Fernandes-Alnemri T, Litwack G;
XX WPI; 1999-130433/11.
XX N-PSDB; AAX04441.
XX Isolated immunophilin FKBP46 nucleic acids - useful for developing
XX products for the study and identification of immunosuppressive
XX agents for treating e.g. transplant and tissue graft patients
XX Claim 1; Fig 1; 21pp; English.
XX This sequence represents the immunophilin FKBP46 from the moth Spodoptera
XX frugiperda. FKBP46 is a novel insect cell nuclear immunophilin, which
XX can be used to study and identify additional immunosuppressive drugs that
XX bind to it. Such immunosuppressive agents can be used in treating
XX transplant and tissue graft patients.
XX Sequence 412 AA;

Query Match 19.6%; Score 311; DB 20; Length 412;
Best Local Similarity 31.0%; Pred. No. 7.3e-17;
Matches 100; Conservative 39; Mismatches 116; Indels 68; Gaps 14;

Qy 3 FMGVAVTP-KNATKVTPEEDSLVHISQSLDCTVKSGESVVLSTVGGAKLVIGTLSQDK 61
Db ||| : | | | : | ||||| : : | | | : | | : || : |
2 fwglmcpnkrytqv---ekpfthisqamdsgndndpcqvmvvdgknflvctliqkgk 58

Qy 62 FPQISFDLVDKEEELSH-SGTANVHIFIGYKSPNTEQD--DTSSDDEDVPEAVPAPAP 118
Db ||| : | | | : | ||||| : | | | : | | : ||| : |
59 llqpldllyfsgdsvalfngkcnvltqylpdefeedledeeeaeaeaeapplv 118

Qy 119 T---AVTANGACAAVYKADTKKAKPAEVKPAEKPSDEED-----ESDDED 164
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
119 aknkrklenandatanlkakpkagknsapaesdddddqlqkflqgedididend 178

Qy 165 ESEDDDSKGMVDSDSDDEE-----EDSEDEEEETPKKPEPINKKRP----- 211
Db || : | : | : | : | : | : | : | : | : | : | : | : |
179 esfkmntsaeegdsdeedddeedddeeddeeeaeapkk-----kkkqpaeeqdstld 234

Qy 212 --NESV--SKTPVSGK-----KAKPAAPASTP-----OKTEKKKGHTATPH 250
Db |||| | | | | | | | | | | | | | | | | | | | | | |
235 tskesvdmsklssqkrlrllkklqqakqpqngvdkpkkeppqkaekkk----- 286

Qy 251 PAKKGKSPVNAVNSPK-SGQGS 272
Db | : | : | : | : | : | : | : | : | : | : | : |
287 peakkeapvekkekkiaagvs 309

```

```

RESULT 15
AAG09171
ID AAG09171 standard; Protein; 203 AA.
XX AC AAG09171;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 6998.
XX Arabidopsis thaliana.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.

```

PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUL-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143542.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144684.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159838.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145132.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161920.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161922.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			

Query Match

Best Local Similarity

Matches

80: Conservative

33: Mismatches

75: Indels

71: Gaps

11:

QY	1	MEFWGAVTPKNAIKVTPEEDSLVHISQASLDCT--VKGESVVLSTVGGAK--LVIGT	56
Db	1	metfwgieikpgkpfkvikqdgfmvhasqvllgdvekvkddetfavvykigddengfmign	60
QY	57	LSQDKFPQISFDLVDFKFEFELSHSGTKANYVHFICYKSPNI--EODDFTSSDDEDDVPRVP	114
Db	61	lsq-kfpqfsidlylghefeishnst--ssvyligrytfdafeideidsdse-----	111
QY	115	APAPTAVTANGAGAAVVVKADTKPKAPAEVKPAEKPESDESEDESEDDDDSEK	174
Db	112	-----ldeymeqgiaalpqrneinpeeddsdsd-----em	141

Search completed: May 2, 2002, 22:03:53
Job time: 6541 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	544	34.2	311	4	US-09-282-305-16	Sequence 16, Appl
2	535	33.7	302	4	US-09-282-305-14	Sequence 14, Appl
3	528	33.2	305	4	US-09-282-305-12	Sequence 12, Appl
4	436	27.4	285	4	US-09-282-305-18	Sequence 18, Appl
5	311	19.6	412	2	US-08-741-134-2	Sequence 2, Appl
6	206	13.0	714	2	US-08-990-114-3	Sequence 3, Appl
7	206	13.0	714	4	US-09-241-333-3	Sequence 3, Appl
8	188	11.8	723	1	US-07-814-964-11	Sequence 11, Appl
9	188	11.8	723	1	US-08-258-442-11	Sequence 11, Appl
10	188	11.8	723	1	US-08-328-809-6	Sequence 6, Appl
11	188	11.8	723	5	PT-US92-11107-11	Sequence 11, Appl
12	180	11.3	1162	2	US-08-728-323A-2	Sequence 2, Appl
13	170	10.7	764	1	US-08-375-300-4	Sequence 4, Appl
14	170	10.7	764	3	US-09-177-431-4	Sequence 4, Appl
15	170	10.7	764	5	PT-US95-16930-4	Sequence 4, Appl
16	170	10.7	1089	1	US-08-375-300-2	Sequence 2, Appl
17	170	10.7	1089	3	US-09-177-431-2	Sequence 2, Appl
18	170	10.7	1089	5	PT-US95-16930-2	Sequence 2, Appl
19	169	10.6	292	4	US-09-538-871-2	Sequence 2, Appl
20	168.5	10.6	739	4	US-09-022-983-2	Sequence 2, Appl
21	168	10.6	411	2	US-08-741-134-6	Sequence 6, Appl
22	167.5	10.5	740	4	US-09-022-983-5	Sequence 5, Appl
23	166.5	10.5	257	4	US-09-538-871-4	Sequence 4, Appl
24	164	10.3	542	1	US-08-701-380-2	Sequence 2, Appl
25	164	10.3	542	3	US-09-032-365A-13	Sequence 13, Appl
26	161.5	10.2	905	2	US-08-574-959A-9	Sequence 9, Appl
27	161.5	10.2	905	4	US-09-357-014-9	Sequence 9, Appl

[illegible]

Db	269	GSASALESHQAKKHEA	285
<hr/>			
Db	235	TSKESVDMSKLSKSQRRLKKLQQOAKQOPQVNGVDPKKEPQOKAEKK	286
QY	251	PAKKGKSPVANOSPK-SGQS	272
Db	287	PEAKKEAPVERKEKKQIAGVS	309
<hr/>			
RESULT	6		
US-08-990-114-3			
; Sequence 3, Application US/08990114			
; Patent No. 5932475			
; GENERAL INFORMATION:			
; APPLICANT: Bandman, Olga			
; APPLICANT: Yue, Henry			
; APPLICANT: Corley, Neil C.			
; APPLICANT: Shah, Purvi			
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN			
; NUMBER OF SEQUENCES: 3			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Incyte Pharmaceuticals, Inc.			
; STREET: 3174 Porter Drive			
; CITY: Palo Alto			
; STATE: CA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: DOS			
; SOFTWARE: FastSeq for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/990,114			
; FILING DATE: Herewith			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Billings, Lucy J.			
; REGISTRATION NUMBER: 36,749			
; REFERENCE/DOCKET NUMBER: PF-0451 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 650-855-0555			
; TELEFAX: 650-845-4166			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 714 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: GenBank			
; CLONE: 128842			
US-08-990-114-3			
<hr/>			
Query Match	13.0%;	Score 206;	DB 2; Length 714;
Best Local Similarity	23.2%;	Pred. No. 1.2e-09;	
Matches	66;	Conservative 31;	Mismatches 81; Indels 106; Gaps
QY	97	EODDETSSDDEOVPE-----AVPAPA-----	117
Db	35	EEEDDSGGSEVVIPQKKGKKATATPAKVVSOTKKVAVPTPAKKAATPAK	94
QY	118	-----PTAVTANGAAVVVKADTKPKAKPAEVKPAE-----EKPESEDESEDEDE	165
Db	95	KAVTPAKAVATPGCKGATQAKALVATPGKGAVTPAKGAKNKNAKKEDESDDDDD	154
QY	166	SEDDDSSEKGMVDDEDD-----SDDDFEFSEDEFEETPKKP	205
Db	155	DDSEDEEE---DEEDEFPPVVKOGKVAAPASEDEDEDEDEDEDEDE---	206
<hr/>			
Db	269	GSASALESHQAKKHEA	285
RESULT	5		
US-08-741-134-2			
; Sequence 2, Application US/08741134			
; Patent No. 5861498			
; GENERAL INFORMATION:			
; APPLICANT: Litwack, Gerald			
; APPLICANT: Alnemri, Emad S.			
; APPLICANT: Fernandes-Alnemri, Teresa			
; TITLE OF INVENTION: IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING			
; TITLE OF INVENTION: AND			
; TITLE OF INVENTION: METHODS OF USING THE SAME			
; NUMBER OF SEQUENCES: 6			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris			
; STREET: One Liberty Place - 46th floor			
; CITY: Philadelphia			
; STATE: Pennsylvania			
; COUNTRY: USA			
; ZIP: 19103			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: Windows 3.11			
; SOFTWARE: Wordperfect for Windows 6.1			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/741,134			
; FILING DATE:			
; CLASSIFICATION: 424			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 60/007,163			
; FILING DATE: 01-NOV-1995			
; CLASSIFICATION: 424			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Deluca, Mark			
; REGISTRATION NUMBER: 33,229			
; REFERENCE/DOCKET NUMBER: TJU-2090			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 215-568-3100			
; TELEFAX: 215-568-3439			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 412 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-741-134-2			
<hr/>			
Query Match	19.6%;	Score 311;	DB 2; Length 412;
Best Local Similarity	31.0%;	Pred. No. 6.8e-19;	
Matches	100;	Conservative 39;	Mismatches 116; Indels 68; Gaps 14;
QY	3	FVGAVVTP-KNATKVTPEEDSLVHISOASLDCTVKSGSVLVSTVGGAKLVIGTLISODK	61
Db	2	FWGLIMEPNKRYTOVV---EKPHISOAMDISTGNDPCQVWVVDGRNFLVCTLOKKG	58
QY	62	FPOISDLVDKFEEELSH-SGTKANVHFIFYKSPNTIQD--DFTSSDDEDVPVAPAPAP	118
Db	59	IIOVPLDLVFKGSDVSFLTNKCIVHLTCYLDPFEEDLEDEEAEEEEEEAPLPVP	118
QY	119	T----AVTANGAACAAVVKADTKPKAKPAEVKPAEBEPSEDEED-----ESDDED	164
Db	119	AKNRKLENANDATAANKKAPDKKGNKSAPAESDSDDDDEQLQKFLDGEDIDTDEND	


```

; FEATURE:
; NAME/KEY: Domain
; LOCATION: 547..620
; OTHER INFORMATION: /label- HMG-box
;
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label- Basic II
;
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label- Mixed Charge
;
US-07-814-964-11

Query Match 11.8%; Score 188; DB 1; Length 723;
Best Local Similarity 23.5%; Pred. No. 4.2e-08;
Matches 76; Conservative 53; Mismatches 110; Indels 84; Gaps 14;

Qy 19 EEDSLVHISQASLDCTVKSGESVVLSTV-GGAKLVIGTISQDKFPOISFDLVDFDKFEL 77
Db 372 EEISSVNFARSG-----GSTRSDFEVLTKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425
Qy 78 SHSGTKANVHFVIGYKSPNIEODDFTSSDDDEVPAPAPAPTAVTANGNAGAAYVVKADTK 137
Db 426 SNMGKDKS-----GYKDV-----DFGSDNENEPDAY-----LARKAEAR 461
Qy 138 PKAKPAEVKPAEKPESDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 187
Db 462 EK-----EEDDDGDSDEESTDEDFKPNENESDVAEYDSDNVEDSDDDSDASGGGG 513
Qy 188 EEDSEDEEEETPKKPEPINKKRNESVSTPTVSGKK--AKPAAPAS-----TPQKT 238
Db 514 DSDGAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 573
Qy 239 EKKKGHTATPHPAKGGK-----SPVNAOSPKSGQS---S 273
Db 574 KRNPGIKVT-EIAKKGEMMKELKDKSKWEDAAAKDKQRYHDEMRYKPEAGGSDNEK 632
Qy 274 GGNKKKPFNSGKQFGGSNNKGS 296
Db 633 GKKSKKRRKTEPSPSKKANTSGS 655

RESULT 9
US-08-258-442-11
; Sequence 11, Application US/08258442
; Patent No. 5670621
; GENERAL INFORMATION:
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pili, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellert, Patti
; APPLICANT: Essigmann, John M.
; APPLICANT: Lippard, Stephen J.
; TITLE OF INVENTION: DNA Structure Specific Recognition
; TITLE OF INVENTION: Protein and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/258,442
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/539,906
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4787AAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 11:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Drosophila melanogaster
; IMMEDIATE SOURCE:
; CLONE: Drosophila SSRP (predicted)
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 458..507
; OTHER INFORMATION: /label- Acidic
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 518..547
; OTHER INFORMATION: /label- Basic I
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 547..620
; OTHER INFORMATION: /label- HMG-box
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 632..649
; OTHER INFORMATION: /label- Basic II
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 657..723
; OTHER INFORMATION: /label- Mixed Charge
;
US-08-258-442-11

```

```

Query Match 11.8%; Score 188; DB 1; Length 723;
Best Local Similarity 23.5%; Pred. No. 4.2e-08;
Matches 76; Conservative 53; Mismatches 110; Indels 84; Gaps 14;

Qy 19 EEDSLVHISQASLDCTVKSGESVVLSTV-GGAKLVIGTISQDKFPOISFDLVDFDKFEL 77
Db 372 EEISSVNFARSG-----GSTRSDFEVLTKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425
Qy 78 SHSGTKANVHFVIGYKSPNIEODDFTSSDDDEVPAPAPAPTAVTANGNAGAAYVVKADTK 137
Db 426 SNMGKDKS-----GYKDV-----DFGSDNENEPDAY-----LARKAEAR 461
Qy 138 PKAKPAEVKPAEKPESDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 187
Db 462 EK-----EEDDDGDSDEESTDEDFKPNENESDVAEYDSDNVEDSDDDSDASGGGG 513
Qy 188 EEDSEDEEEETPKKPEPINKKRNESVSTPTVSGKK--AKPAAPAS-----TPQKT 238
Db 514 DSDGAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 573
Qy 239 EKKKGHTATPHPAKGGK-----SPVNAOSPKSGQS---S 273
Db 574 KRNPGIKVT-EIAKKGEMMKELKDKSKWEDAAAKDKQRYHDEMRYKPEAGGSDNEK 632
Qy 274 GGNKKKPFNSGKQFGGSNNKGS 296
Db 633 GKKSKKRRKTEPSPSKKANTSGS 655

```

Db 633 GCKSKKRKTEPSKANTSGS 655

RESULT 10

US-08-328-809-6
; Sequence 6, Application US/08328809
; Patent No. 5705334
; GENERAL INFORMATION:
; APPLICANT: Lippard, Stephen J.
; APPLICANT: Essigmann, John M.
; APPLICANT: Donahue, Brian A.
; APPLICANT: Toney, Jeffrey H.
; APPLICANT: Bruhn, Suzanne L.
; APPLICANT: Pil, Pieter M.
; APPLICANT: Brown, Steven
; APPLICANT: Kellett, Patti
; TITLE OF INVENTION: Uses For DNA Structure-Specific
; TITLE OF INVENTION: Recognition Proteins

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patent Administrator, Testa, Hurwitz & Thibault

STREET: 53 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/08/328,809

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fenton, Gillian M.

REGISTRATION NUMBER: 36,508

REFERENCE/DOCKET NUMBER: MIT-023 (5473/24)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-248-7000

TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 723 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Drosophila melanogaster

IMMEDIATE SOURCE:

CLONE: Drosophila SSRP (predicted)

FEATURE:

NAME/KEY: Domain

LOCATION: 458..507

OTHER INFORMATION: /label= Acidic

FEATURE:

NAME/KEY: Domain

LOCATION: 518..547

OTHER INFORMATION: /label= Basic I

FEATURE:

NAME/KEY: Domain

LOCATION: 547..620

OTHER INFORMATION: /label= HMG

FEATURE:

NAME/KEY: Domain

LOCATION: 632..649

OTHER INFORMATION: /label= Basic II

FEATURE:

NAME/KEY: Domain

LOCATION: 657..723

OTHER INFORMATION: /label= Mixed Charge

US-08-328-809-6

Query Match 11.8%; Score 188; DB 1; Length 723;
Best Local Similarity 23.5%; Pred. No. 4.2e-08;
Matches 76; Conservative 53; Mismatches 110; Indels 84; Gaps 14;

QY 19 EEDSLVHISOASIDCTVKSGESVLSVTV-GGAKLVIGTLSDQKPPQISFDLVFDKEPEL 77

Db 372 EETSSVNFARSG-----GSTRSPDFEVLKNGTVHIFSSIEKEEYAKL-FDYITQKKLHV 425

QY 78 SHSGTKANVHFICYKSPNIEQDDFTSSDDEDPEAVPAPAPTAVTANGAGAAVVAADTK 137

Db 426 SNMKDKS-----GYKDY-----DFGSDNENEPDAY-----LARIKAEAR 461

QY 138 PRAKPAEVKPAEKPESEDESDDESEDDDDSEKGMVD---EDDSDDE----- 187

Db 462 EK-----EEDDDGDSDESTDDEDFPNESESVAEYDSNVESDDSDSDASGGGG 513

QY 188 EFDSEDEEEETPKPEPINKKRPNESVSKTPVSGKK--AKPAAPAS-----TPQKT 238

Db 514 DSDGAKKKKKSEKKEKKHKEKERTKKPSKKKKDGKPKRATTAFMLWLDNDFRESI 573

QY 239 EKKKGCHTATHPAKKGGK-----SPVNAQSPKSGGQS---S 273

Db 574 KRENGIKVT-ETAKKGEMKELDKSKWEDAAAKQRYHDEMKNYKPEAGGDSDNK 632

QY 274 GGNKKKPFNSGKQFGGSNNKGS 296

Db 633 GKSKKRKTEPSKANTSGS 655

RESULT 11

PCT-US92-11107-11

; Sequence 11, Application PC/TUS92111107

; GENERAL INFORMATION:

; APPLICANT: Donahue, Brian A.

; APPLICANT: Toney, Jeffrey H.

; APPLICANT: Bruhn, Suzanne L.

; APPLICANT: Pil, Pieter M.

; APPLICANT: Brown, Steven

; APPLICANT: Kellett, Patti

; APPLICANT: Essigmann, John M.

; APPLICANT: Lippard, Stephen J.

; TITLE OF INVENTION: DNA Structure Specific Recognition

; TITLE OF INVENTION: Protein and Uses Therefor

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: 2 Militia Drive

; CITY: Lexington

; STATE: MA

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/11107

; FILING DATE: 19921218

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/539,906

; FILING DATE: 18-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: MIT-4787AAA

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Drosophila melanogaster
IMMEDIATE SOURCE: Drosophila SSRP (predicted)
FEATURE:
NAME/KEY: Domain
LOCATION: 458..507
OTHER INFORMATION: /label- Acidic
FEATURE:
NAME/KEY: Domain
LOCATION: 518..547
OTHER INFORMATION: /label- Basic I
FEATURE:
NAME/KEY: Domain
LOCATION: 547..620
OTHER INFORMATION: /label- HMG-box
FEATURE:
NAME/KEY: Domain
LOCATION: 632..649
OTHER INFORMATION: /label- Basic II
FEATURE:
NAME/KEY: Domain
LOCATION: 657..723
OTHER INFORMATION: /label- Mixed Charge
PCT-US92-11107-11

Query Match 11.8%; Score 188; DB 5; Length 723;
Best Local Similarity 23.5%; Pred. No. 4.2e-08;
Matches 76; Conservative 53; Mismatches 110; Indels 84; Gaps 14;
QY 19 EEDSLVHISQASLDCTVKSGESVLSVTV-GGAKLVIGTILSDKFFPISDLDVDFEFEL 77
Db 372 EEISVNFARSG-----GSTRSDFEVLKNGTVHFISIEKEEYAKL-FDYITQKLVH 425
QY 78 SHSGTKANVHFIQYKSPNIEODDFTSSDDEVDPEAVPAPAPTAVTANGNAGAIVVAKDTK 137
Db 426 SNMKDKS---GYKDV-----DFGSDNENEPDAY-----LARKAEAR 461
QY 138 PKAKPAEVKPAEKPEDEDEDEDEDDSEKGMVDV-----EDSDDE----- 187
Db 462 EK-----EEDDDGDSDEESTDEDFKPNENESDVAEYDSNVESDSDSDSDASGGG 513
QY 188 EEDSEDEEEETPKKPEPKKPNESVSKTPVSGKK--AKPAAAPAS-----TPQKT 238
Db 514 DSGAKKKKKKKSEKK 573
QY 239 EKKKGHTATPHPAKGGK-----SPVNAQSPKSGGQS---S 273
Db 574 KRENPGIKVT-EIAKGGEMKELDKSKWEDAAAKDKQRYHDEMRYKPEAGSDSNEK 632
QY 274 GGNKKKPFNSGGKQFGSGNNKGS 296
Db 633 GKSCKKRRKTEPSPSKKANTSGS 655
RESULT 12
US-08-728-323A-2
Sequence 2, Application US/08728323A
Patent No. 5948676
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-728-323A-2
Query Match 11.3%; Score 180; DB 2; Length 1162;
Best Local Similarity 27.8%; Pred. No. 3.7e-07;
Matches 49; Conservative 18; Mismatches 55; Indels 54; Gaps 3;
QY 103 SSDEVDPEAVPAPAPTAV-----TANGNAG 128
Db 248 SSSDGDTPPRQPTSPISGSSSPSGSGMDDTAMLVLLAEAEASKNKEKSENNOAC 307
QY 129 -----AAVWADTKPAKPAVPAEKPEDEE-----DESDDESE 168
Db 308 EDNGDNEISKESQVDRDNKDEEEQETDERDEDEDEDEDEDEDEDEDEDEDEDE 367
QY 169 DDBSEKGMVDDESDDE 224
Db 368 DDE 423
RESULT 13
US-08-375-300-4
Sequence 4, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

RESULT 14
US-09-177-431-4
: Sequence 4, Application US/09177431
: Patent No. 6071700
: GENERAL INFORMATION:
: APPLICANT: He, Feng
: APPLICANT: Jacobson, Allan S.
: TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
: TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: Fastseq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/177.431
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/955,472

```

RESULT 15
PCT-US95-16930-4
; Sequence 4, Application PC/TUS9516930
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
; TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
; TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY
; TITLE OF INVENTION: FUNCTION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16930
; FILING DATE: 27-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375,300
; FILING DATE: 20-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:07:30 ; Search time 205.36 Seconds
(without alignments)
142.712 Million cell updates/sec

Title: US-09-645-337-8
Perfect score: 1589
Sequence: 1 MEFWGVAVTPKNATKVTPEE.....KQFGSNNKSGNKCKGKGRA 305

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589	100.0	305	2 T52287	probable histone d
2	543.5	34.2	307	2 T04141	histone deacetylase
3	450.5	28.4	296	2 T48401	histone deacetylase
4	311	19.6	412	2 A55320	immunophilin FKBP4
5	243	15.3	357	2 JC4090	FK506-binding 39k
6	230	14.5	798	2 T33022	hypothetical prote
7	215.5	13.6	694	1 DNCHNL	nucleolin - chicke
8	211.5	13.3	712	2 JH0148	nucleolin - rat
9	207.5	13.1	611	2 T06458	nucleolin homolog
10	206.5	13.0	707	1 DNMS	nucleolin - mouse
11	206	13.0	713	2 A27441	nucleolin - human
12	205	12.9	707	2 A35804	heat-shock protein
13	200.5	12.6	281	2 JC4295	nucleolin homolog
14	196.5	12.4	635	2 T09648	hypothetical prote
15	192.5	12.1	971	2 T19431	hypothetical prote
16	188	11.8	723	2 S33688	nucleolin - cytoplas
17	180	11.8	734	2 B42680	single-strand DNA/
18	186	11.7	723	2 A48217	hypothetical prote
19	185	11.6	751	2 T24541	nucleophosmin - ch
20	183.5	11.5	294	1 DNCHFM	probable peptide-a
21	182.5	11.5	872	2 T18861	transcription fact
22	182.5	11.5	1187	2 T46637	protein F32E10.6 l
23	180.5	11.4	374	2 C88734	probable immediate
24	180	11.3	1300	2 T03166	transcription fact
25	178.5	11.2	517	2 S20248	ATP-dept. acyl-CoA
26	177.5	11.2	885	2 G71608	suppressor of sabl
27	177.5	11.2	1473	2 T13855	neurofilament prot
28	177	11.1	644	2 S55395	nucleolin - Africa
29	177	11.1	651	2 S18874	

ALIGNMENTS

RESULT 1

T52287

Probable histone deacetylase (EC 3.5.1.-) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000

C:Accession: T52287

R:Dangl, M.; Haas, H.; Loidl, P.

submitted to the EMBL Data Library, January 1998

A:Description: Arabidopsis thaliana complete cDNA-sequence homologous to Zea mays HD2

A:Reference number: 226016

A:Accession: T52287

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-305 <DAN>

A:Cross-references: EMBL:AF044914; PIDN:AAC02539.1

A:Experimental source: cultivar Columbia

C:Genetics:

A:Gene: HD2

C:Keywords: hydrolase

Query Match	100.0%	Score 1589;	DB 2;	Length 305;
Best Local Similarity	100.0%	Pred. No. 6.6e-78;		
Matches 305;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	MEFWGVAVTPKNATKVTPEEDSLVHISQASLDC	TVKSGESVLSVTVGAKLVIGT	LSQD 60	
DB 1	MEFWGVAVTPKNATKVTPEEDSLVHISQASLDC	TVKSGESVLSVTVGAKLVIGT	LSQD 60	
QY 61	KFPOISFDLVDFKFEFELSHSGTKANVHF	IGYKSPNIEQDDFTSSDDE	YPEAVPAPAPTA 120	
DB 61	KFPOISFDLVDFKFEFELSHSGTKANVHF	IGYKSPNIEQDDFTSSDDE	YPEAVPAPAPTA 120	
QY 121	VTANGNAGAAVVKADTKPKAKPAEVKPAE	KPEKPEDESDDEDEDEDE	DDSEDDSEDDSEDDSE 180	
DB 121	VTANGNAGAAVVKADTKPKAKPAEVKPAE	KPEKPEDESDDESDDESD	DEDESDDESDDESDDE 180	
QY 181	DDSDDDDEDEDEDEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDE 240	
DB 181	DDSDDDDEDEDEDEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDE 240	
QY 241	KKGGHTATPHAKKGGKSPVNAQSPKSGQ	SGGNNKPKFPNSGKFGG	SGNNKSGK 300	
DB 241	KKGGHTATPHAKKGGKSPVNAQSPKSGQ	SGGNNKPKFPNSGKFGG	SGNNKSGK 300	
QY 301	GKGRA 305			
DB 301	GKGRA 305			
RESULT 2				
T04141				

histone deacetylase (EC 3.5.1.-) HD2-p39, nucleolar - maize
 C:Species: Zea mays (maize)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: T04141
 R:Lusser, A.; Brosch, G.; Loidl, A.; Haas, H.; Loidl, P.
 Science 277, 88-91, 1997
 A>Title: Identification of maize histone deacetylase HD2 as an acidic nucleolar phosphoprotein
 A:Reference number: 215237; MUID:97349336
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-307 <IUS>
 A:Cross-references: EMBL:U02815; NID:g2257755; PIDN:AAB63262.1; PID:g2257756
 A:Experimental source: strain Cuzco 251
 C:Keywords: hydrolase

Query Match 34.2%; Score 543.5; DB 2; Length 307;
 Best Local Similarity 44.8%; Pred. No. 2.4e-22;
 Matches 130; Conservative 45; Mismatches 96; Indels 19; Gaps 11;
 QY 1 MEFGVAVTPKNAKVTPEEDSLVHSQSASLDCTVKSSESWSLVTGGAKLVTGLSD 60
 DQ 1 MEFGVAVTPKNAKVTPEEDSLVHSQSASLDCTVKSSESWSLVTGGAKLVTGLSD 60
 QY 61 KFOISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDFTSSDDEDPVPAAPAPTA 120
 DQ 60 KNPIHQDLIFDKFEFELSHSGTKANVHFICYKSPNIEQDFTSSDDEDPVPAAPAPTA 120
 QY 121 VTANGNAGAAVAVKADTKPAKPAEVKPAEKPESDESDDESDDESD--DDSEKGM 178
 DQ 117 VKENKADEKQKQSKAEVAAAPSPDSKSKDDSDSDDESDDESDDESDDESDGLSS 176
 QY 179 DEEDSDDESDDESDDESDDESD--TPKKPEPINKKRPNE-SVSKTPVSGKAKKAPAAAPASTP 235
 DQ 177 ERGDDSDSDDESDDESDDESDDESD--TPKKPE-VGKKRPAESSVLKTPSLDKKAK-VATPSS-- 232
 QY 236 OKTEKKKGG--HTATPHPAKKGKSPVNAQSPKS--GGSSGCGNNKKP 281
 DQ 233 QKTGGKGAHVATPHPAK--GKTIYNNDKSVKSPKSPKSGGVPCKP 280
 RESULT 3
 T48401
 histone deacetylase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F17C15.160
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48401
 R:Bevan, M.; Pohl, T.; Weizengger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: 224492
 A:Accession: T48401
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-296 <BEV>
 A:Cross-references: EMBL:AL162506
 A:Experimental source: cultivar Columbia; BAC clone F17C15
 C:Genetics:
 A:Map position: 5
 A:Introns: 5/1; 28/3; 93/3; 116/3; 164/3; 195/3; 227/1
 A>Note: F17C15.160

Query Match 28.4%; Score 450.5; DB 2; Length 296;
 Best Local Similarity 38.3%; Pred. No. 2e-17;
 Matches 118; Conservative 39; Mismatches 94; Indels 57; Gaps 12;
 QY 1 MEFGVAVTPKNAKVTPEEDSLVHSQSASLDCTVKSSESWSLVTGGAKLVTGLSD 53
 DQ 1 MEFGVAVTPKNAKVTPEEDSLVHSQSASLDCTVKSSESWSLVTGGAKLVTGLSD 53
 QY 54 IGTLSQDKFPQISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDFTSSDDESD-- 108

Db 55 IGTLSQDKFPQISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDFTSSDDESD 114
 QY 109 VPEAVPAPAPTAVTANGNAGAAVAVKADTKPAKPAEVKPAEKPESDESDDESD 168
 Db 115 YPPAAKPSAAKQVNFQ-----LPNIDVAVK-----QDDADGSESDSDSDSDS 160
 QY 169 DDISEKGM-----DVIDRSDSDDESDDESDDESDDESDDESDDESDDESD 224
 Db 161 GDEEEKVTAESDSESDSDSDDESDS-----SEETPKKPEEPKPSAEPNSKNPASNKK 217
 QY 225 AKPAAAPASTPQKTEKKKGG--HTATPHPAKKGKSPVNAQSPKSGGSSGCGNNKK--P 281
 Db 218 AK-----FVTQKTDISKKPHVHATPHPSKQAGKN-----SGGSGTGTSTKQQTTP 263
 QY 282 ENSGKQFG 289
 Db 264 KSAG-AFG 270
 RESULT 4
 A55320
 immunophilin FKBP46 - fall armyworm
 C:Species: Spodoptera frugiperda (fall armyworm)
 C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 02-Sep-2000
 C:Accession: A55320
 R:Alnemri, E.S.; Fernandes-Alnemri, T.; Pomeroy, K.; Robertson, N.M.; Dudley, K.; D
 J. Biol. Chem. 269, 30828-30834, 1994
 A>Title: FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a protein-k
 A:Reference number: A55320; MUID:95074110
 A:Accession: A55320
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-412 <ALN>
 A:Cross-references: GB:U15036; NID:g595844; PID:g595845
 A:Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomeras
 C:Keywords: nucleus; phosphoprotein
 F:324-371/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>
 Query Match 19.6%; Score 311; DB 2; Length 412;
 Best Local Similarity 31.0%; Pred. No. 7.1e-10;
 Matches 100; Conservative 39; Mismatches 116; Indels 68; Gaps 14;
 QY 3 FWGAVVTP-KNATKVTPEEDSLVHSQSASLDCTVKSSESWSLVTGGAKLVTGLSD 61
 Db 2 FWGLMEPNKRTQVY---EKPEHISQAAMDISTGNDPCQVMVYVDGKFLVCTLQCK 58
 QY 62 FQOISDLVDFKFEFELSH-SGTKANVHFICYKSPNIEQD--DFTSSDDESDDESDDESD 118
 Db 59 IIQVPLDLYFKSGDSVSLTNGKCNVHLTGYLDPFEDEDEDEDEDEDEDEDEDE 118
 QY 119 T---AVTANGNAGAAVAVKADTKPAKPAEVKPAEKPESDESD-----ESDDED 164
 Db 119 ANKKRLENANDATANKKAKPKKACKNAPAAESDSDSDSDDESDDESDDESD 178
 QY 165 ESEEDSDSEKGMVDSDSDSDDESD-----EDSEDEDEDEDETPKKPEPINKKRP 211
 Db 179 ESFKMNTSAEGDSDSDSDDESDDESDDESDDESDDESDDESDDESDDESD 234
 QY 212 --NESV--SKTPVSGK-----KAKPAAAPASTP-----OKTEKKKGGHTATPH 250
 Db 235 TSKEYSYDMSKSKSQRRLKRLKQQAQKQOQVGVNDKPKKEEPQKAEKKK----- 286
 QY 251 PAKKGGKSPVNAQSPK-SGGS 272
 Db 287 PEAKKEAPVEKKEKKQIAGGS 309
 RESULT 5
 JC4090
 FK506-binding 39k protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster

C:Date: 23-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jul-2000

C:Accession: JC4090
 R:Theopold, U.; Zotto, L.D.; Hultmark, D.
 Gene 156, 247-251, 1995
 A:Title: FRBP39, a Drosophila member of a family of proteins that bind the Immunosuppression
 A:Reference number: JC4090; MUID:95278752
 A:Accession: JC4090
 A:Molecule type: mRNA
 A:Residues: 1-357 <HE>
 A:Cross-references: EMBL:Z46894; NID:g600423; PIDN:CAA86996.1; PID:g600424
 C:Genetics:
 A:Gene: fkbp39
 A:Cross-references: FlyBase:FBgn0013269
 A:Superfamily: yeast peptidylprolyl isomerase FPR3; BKBP-type peptidylprolyl isomerase h
 F:89-99/Region: acidic
 F:252-357/Region: FK506 binding #status predicted
 F:269-316/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>

Query Match 15.38; Score 243; DB 2; Length 357;
 Best Local Similarity 26.98; Pred. No. 2.5e-06;
 Matches 95; Conservative 49; Mismatches 109; Indels 100; Gaps 18;

QY 3 FWGVAVTPKNAKVTTPEDSLVHLSQASLDCTVKSGESVLSVTVGCAKLVIGTLSQDKF 62
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 4 FWGLNMRPER-KVSQTIKSFHSGVALD-----KGEAKLYLAAEQEYIVATVTK-AI 56
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 63 POISFDLVDFK-EFELSHSGTKANVHFIGYKSPNIEQDDFTSSDEDVPEAVPAPTAV 121
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 57 PQVALDLNFSKGDRIIMFYTAGDSVSLGY-----LHDIDSEDEDDQ-----M 101
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 122 TANGNACAANVADTKPKAKPAEVPKPAEKPESDEPESDEDES-----E 167
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 102 TIENLLNSKALKNSKKS-----EDDENESGEDEEDTDDDSQIIEYESFLENLEE 154
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 168 EDDSEKGMVDYED-DSDDDEEDSDREE-EETPKKPEPKPKRPNESVSKTPVSGKK 224
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 155 EDDO-----DVDENESEGEDEQSDSEAEEOQPR--KVAKLSQASAKK---SGKE 204
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 225 AKPAAPASTPQTEKKKGHTATPHPAKKGKSPVNAOSPKS----- 268
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 205 QNGVAKKEAKQKQKKE-----KPEAKKEQPKAKPAKQKQPKASKDPRITITGGVKIQDV 259
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 269 ---GQSSGG-----NNKKKPNFNS---CKQF-----GGSNNKSGNK 299
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 260 VGKGEAKQGRVYVYIGRLQSNKTKFDSLLKGPKFKALGGGEVKGWDVG 312
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 6
 T33022
 hypothetical protein K07H8.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33022
 R:Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.
 submitted to the EMBL Data Library, February 1998
 A:Description: The sequence of C. elegans cosmid K07H8.
 A:Reference number: 221264
 A:Accession: T33022
 A:Status: preliminary;
 A:Molecule type: DNA
 A:Residues: 1-798 <FUL>
 A:Cross-references: EMBL:AF047659; PIDN:AAC04430.1; GSPDB:GN00022; CESP:K07H8.10
 A:Experimental source: strain Bristol N2; clone K07H8
 C:Genetics:
 A:Gene: CESP:K07H8.10
 A:Map position: 4
 A:Introns: 205/1; 308/3; 630/1; 773/3

Query Match 14.58; Score 230; DB 2; Length 798;
 Best Local Similarity 31.78; Pred. No. 2.7e-05;
 Matches 72; Conservative 32; Mismatches 83; Indels 40; Gaps 9;

QY 48 GGAKLVIGTLSQDKFPQISFDLVDFKFEFELSHSGTKANVHFIGYKSPNIEQDDFTSSDDE 107
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 181 GGSRVGF-----SDKRQYDSGDGDEEEVPPKKKLTATGTTFKQTKP-VEDD--SGEDEE 233
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 108 DVEAVPAPAPTAVTANGNACAANVADTKPKAKPAEVPKPAEKPESDEDESDESE 167
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 234 DEESDEEPQF-----PKTKAKSAIATAAE---DSDDDEEDDEDEEE 273
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 168 EDDDS-----EKGMDVDEDDSDDEEE-----DSEDEEEETPKKPEPKPKRPNESVSKTPV 220
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 274 SDDAPQVKKSAVEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDETEPVAKTPAVKSSLSKSD 333
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 QY 221 SGKAKPAAAPASTPQTEKKKGHTATPHPAKKGKSPVNAOSPK 267
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 DB 334 STRGKQVNLKSVATPTTVDKKA---PATPHPAKTSKNA---AGDVPK 374
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 RESULT 7
 DNCHNL
 nucleolin - chicken
 N:Alternate names: nucleolar protein C23
 C:Species: Gallus gallus (chicken)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
 C:Accession: S08414; S10766; A32725; I50397; B30099
 R:Maridor, G.; Nigg, E.A.
 Nucleic Acids Res. 18, 1286, 1990
 A:Title: cDNA sequences of chicken nucleolin/C23 and N038/B23, two major nucleolar pr
 A:Reference number: S08414; MUID:90206792
 A:Accession: S08414
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-694 <MAR>
 A:Cross-references: EMBL:X17199; NID:g63710; PIDN:CAA35060.1; PID:g63711
 R:Maridor, G.; Krek, W.; Nigg, E.A.
 Biochim. Biophys. Acta 1049, 126-133, 1990
 A:Title: Structure and developmental expression of chicken nucleolin and N038: coordi
 A:Reference number: S10766; MUID:90304215
 A:Accession: S10766
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-694 <MAR2>
 R:Peter, M.; Nakagawa, J.; Doree, M.; Labbe, J.C.; Nigg, E.A.
 Cell 60, 791-801, 1990
 A:Title: Identification of major nucleolar proteins as candidate mitotic substrates o
 A:Reference number: A32725; MUID:90182668
 A:Accession: A32725
 A:Molecule type: protein
 A:Residues: 56-62; 63-109; 197-214 <PET>
 R:Borer, R.A.; Lehner, C.F.; Eppenberger, H.M.; Nigg, E.A.
 Cell 56, 379-390, 1989
 A:Title: Major Nucleolar Proteins Shuttle between Nucleus and Cytoplasm.
 A:Reference number: I50397; MUID:89119560
 A:Accession: I50397
 A:Status: translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 407-418; 'R', 420-519; 'T', 521-694 <BOR>
 A:Cross-references: GB:M21791; NID:g212411; PIDN:AAA48983.1; PID:g212412
 C:Comment: Phosphorylation of this protein by cdc2 kinase may contribute to the reorg
 C:Superfamily: nucleolin; ribonucleoprotein repeat homology
 C:Keywords: DNA binding; duplication; nucleolus; nucleus; phosphoprotein; RNA binding
 F:1-247/Domain: acidic <ACI>
 F:54-91/Region: 7-residue repeats (T-P-A-K-K-A-[A/V])
 F:254-262/Region: nuclear location signal
 F:267-275/Region: nuclear location signal
 F:282-347/Domain: ribonucleoprotein repeat homology <RRM1>
 F:283-361/Domain: RNA binding #status predicted <RNA1>
 F:372-435/Domain: ribonucleoprotein repeat homology <RRM2>
 F:373-448/Domain: RNA binding #status predicted <RNA2>
 F:462-525/Domain: ribonucleoprotein repeat homology <RRM3>
 F:463-538/Domain: RNA binding #status predicted <RNA3>
 F:554-618/Domain: ribonucleoprotein repeat homology <RRM4>
 F:555-631/Domain: RNA binding #status predicted <RNA4>

F:632-694/Domain: glycine/arginine-rich <RRM>
F:56,63,70,77,85/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predi

Query Match 13.6%; Score 215.5; DB 1; Length 694;
Best Local Similarity 28.1%; Pred. No. 0.00014;
Matches 85; Conservative 32; Mismatches 102; Indels 83; Gaps 12;
QY 6 VAVTPKNAK-----VTPEEDSLVHISOASLDCTVKSGSVL-----SVT----- 46
Db 4 LAKTPKQMKQKKWAPPPKVESESEESDLESSGEVVPVPPKQKQKAAVTPAKKAAT 63
QY 47 -----VGGAKLVIGTISQDKFPQISDFLVDFKESLHSGTKANVHFICYKSP 94
Db 64 PAKKAATPAKKAATPAKKAAT-----PAKKAAPSPKAAVVGKAKGNKAKKEESE 117
QY 95 NIEODDFTSSDDDDVPAVPAPAPATATANGNAGAAVVKADTPKAKPAVVKPAEKPEPS 154
Db 118 EEDDDDE 171
QY 155 DEEDDE 187
Db 172 EEEDE 231
QY 188 EEDDE 235
Db 232 EDE 290
QY 236 QK 237
Db 291 TK 292

RESULT 8
JH0148
nucleolin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C:Accession: JH0148; A24088; I63130
R:Bourbon, H.M.; Amalric, F.
Gene 88, 187-196, 1990
A:Title: Nucleolin gene organization in rodents: highly conserved sequences within three
A:Reference number: JH0148; MUID:90269607
A:Accession: JH0148
A:Molecule type: DNA
A:Residues: 1-712 <BOU>
A:Note: the authors translated the initiation codon GTG for residue 1 as Val
R:Lischwe, M.A.; Cook, R.G.; Ahn, Y.S.; Yeoman, L.C.; Busch, H.
Biochemistry 24, 6025-6028, 1985
A:Title: Clustering of glycine and Ng,Ng-dimethylarginine in nucleolar protein C23.
A:Reference number: A24088; MUID:86104094
A:Accession: A24088
A:Molecule type: protein
A:Residues: 651-703 <LIS>
R:Bourbon, H.
Gene 68, 73-84, 1988
A:Title: Sequence and structure of the nucleolin promoter in rodents: Characterization of
A:Reference number: I48118; MUID:89121496
A:Accession: I63130
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 'MV', 2-44 <RES>
A:Cross-references: GB:M22090; NID:g205793; PIDN:AAA41733.1; PID:g205794
C:Comment: This protein is the major nucleolar-specific protein in eukaryotic exponentia
C:Genetics:
A:Gene: nuc
A:Start codon: GTG
A:Introns: 5/3; 44/3; 211/1; 274/1; 303/1; 350/2; 392/1; 433/2; 485/1; 526/2; 571/1; 613/1
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleus
F:311-376/Domain: ribonucleoprotein repeat homology <RRM1>
F:397-459/Domain: ribonucleoprotein repeat homology <RRM2>
F:489-552/Domain: ribonucleoprotein repeat homology <RRM3>

F:575-639/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 13.3%; Score 211.5; DB 2; Length 712;
Best Local Similarity 29.4%; Pred. No. 0.00023;
Matches 57; Conservative 27; Mismatches 55; Indels 55; Gaps 7;
QY 97 EODDFTSSDDDDVPAVPAPAPATATANGNAGAAVVKADTPKAKPAVVKPA-----E 149
Db 150 EEDDDSDDE 190
QY 150 EKPEDE 186
Db 191 EEDDDDDDDDE 250
QY 187 EEDSDSEDEEETPKPEPINKRPNESVSKTPVSGKAKKAPAAAPASTPQ-KTEKKKGHI 245
Db 251 DEEDEDEDEDE-----DOEDEDEDEDEPEVKAAP--GRRKKMTKQKEAPEAKKQKEGSE 305
QY 246 TATPHPAKKGKSP 259
Db 306 PTPFNLFIGNLP 319

RESULT 9
T06458
nucleolin homolog - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C:Accession: T06458
R:Tong, C.G.; Hsieh, H.L.; Blumenthal, S.; Reichler, S.; Balk, J.; Roux, S.J.
submitted to the EMBL Data Library, August 1995
A:Description: Molecular cloning and characterization of a cDNA encoding a nucleolin-
A:Reference number: Z15692
A:Accession: T06458
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-611 <TON>
A:Cross-references: EMBL:L43510; NID:g940287; PIDN:AAA74208.1; PID:g940288
A:Experimental source: cv. Alaska
C:Superfamily: nucleolin; ribonucleoprotein repeat homology

Query Match 13.1%; Score 207.5; DB 2; Length 611;
Best Local Similarity 30.9%; Pred. No. 0.00033;
Matches 64; Conservative 24; Mismatches 54; Indels 65; Gaps 8;
QY 97 EODDFTSSDDDDVPAVP-----APATAVTANGNAGAAVVKADTK-PKAK--PAVVKP- 147
Db 153 ESEDESEDEKPAKPAVPAKNGSAPAKKAASSDEDEDEDEDEDEDEDEDEDEDEDE 212
QY 148 -AEEKPES-----DEED-----ESDDEDESEEDDS----- 172
Db 213 AAKKKAESSDSSDSDDEDETKPTVTAVSKSAAVAKKTDEDDDESSDSSDSSDK 272
QY 173 -----EKGMVDVDDDDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 215
Db 273 TSKPSNGKKPVTVSKKEDKMNVDKSSDSDSDSDSDSDSDSDSDSDSDSDSDSD 332
QY 216 SKTPVSGKAKPAAPASTPQKTEKK 242
Db 333 D-----ADKSSKAPATPATPSENGGSK 355

RESULT 10
DNMS
nucleolin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
C:Accession: A29958; A40769; A56240; I84688
R:Bourbon, H.M.; Lapeyre, B.; Amalric, F.
J. Mol. Biol. 200, 627-636, 1988
A:Title: Structure of the mouse nucleolin gene. The complete sequence reveals that ea

C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 23-Jul-1999
C:Accession: A27441; A24808; A24070
R:Lapeyre, B.; Bourbon, H.; Amalric, F.
Proc. Natl. Acad. Sci. U.S.A. 84, 1472-1476, 1987
A>Title: Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual
A:Reference number: A27441; MUID:87175501
A:Molecule type: mRNA
A:Residues: 1-713 <LAP>
R:Lapeyre, B.; Amalric, F.; Ghaffari, S.H.; Venkatarama Rao, S.V.; Dumar, T.S.; Oliso
J. Biol. Chem. 261, 9167-9173, 1986
A>Title: Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing reg
A:Reference number: A24808; MUID:86250857
A:Accession: A24808
A:Molecule type: mRNA
A:Residues: 522-541, 'G', 543-544, 'R' 546-713 <LA2>
A:Cross-references: GB:M13757; NID:g191004; PIDN:AAA36961.1; PID:g387047
R:Lapeyre, B.; Calzergues-Ferrier, M.; Bouche, G.; Amalric, F.
Nucleic Acids Res. 13, 5805-5816, 1985
A>Title: Cloning of cDNA encoding a 100 kDa nucleolar protein (nucleoline) of Chinese
A:Reference number: A24070; MUID:85297777
A:Accession: A24070
A:Molecule type: mRNA
A:Residues: 179-238 <LA3>
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: DNA binding; nucleus; phosphoprotein
F:308-373/Domain: ribonucleoprotein repeat homology <RRM1>
F:394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F:486-549/Domain: ribonucleoprotein repeat homology <RRM3>
F:572-636/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 13.0%; Score 206; DB 2: Length 713;
Best Local Similarity 23.2%; Pred.No.0.00046;
Matches 66; Conservative 31; Mismatches 81; Indels 106; Gaps 8;

QY 97 EQDDFTSSDDEDVPE-----AVPAPA----- 117
||:||:||:||:
Db 34 EEEDSSGEVWIPQKGGKATATPAKKVVVSOTTKVAVTPAKKAATVPKKAATAATPAK 93

QY 118 -----PTAVTANGNAAGVAAVDTKPKAKPAEVKPAE-----EKPSDEDESDDEDE 165
||:||:||:||:
Db 94 KAVTPAKAVATPGKKGATQAKALVATPGKGAVTPAKGAKGNKAKKEDSDEDEDDDDDE 153

QY 166 SEEDDSKEGMVDVEDD-----SDDEDESEDESEETPKKPPEP 205
||:||:||:||:
Db 154 DSDSEDEE---DEEDEFFPPVVGKGVKGVAAAAAPASEDEDEDEEEEEEDEE----EE 205

QY 206 INKRPNESYSKTVPVSGKAKAAPAST-----POKTE 239
||:||:||:||:
Db 206 EDDSEEEAEITPAKCKKAPAKVVPVKAKNAVEEDDDDEDEDEDEDEDEDEDEDEE 265

QY 240 KKKGCHTATPHPAKKGKSPVNANQSPKSGSQSGNNKKPFN 283
||:||:||:||:
Db 266 EEEEEPVPKAPGR-KKEMTKQEVPKAKQKVGESESTTFN 308

RESULT 12
A35804
nucleolin - human
N:Alternate names: phosphoprotein pp100; protein B50; protein C23
C:Species: Homo sapiens (man)
C>Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 23-Jul-1999
C:Accession: A35804; S04631; A48138; A55996
R:Srivastava, M.; McBride, O.W.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.
J. Biol. Chem. 265, 14922-14931, 1990
A>Title: Genomic organization and chromosomal localization of the human nucleolin gen
A:Reference number: A35804; MUID:90368666
A:Accession: A35804
A:Molecule type: DNA
A:Residues: 1-707 <SRI>
A:Cross-references: GB:M60858; GB:J05584; NID:g189305; PIDN:AAA59954.1; PID:g189306
R:Srivastava, M.; Fleming, P.J.; Pollard, H.B.; Burns, A.L.

```
A:Residues: 1-281 <DBM>  
A:Cross-references: GB:I39778; NID:g662358; PION:AAA90510.1; PID:g662359  
C:Comment: This protein has roles in growth, development and thermo-tolerance.  
C:Keywords: heat shock; stress-induced protein
```

```
Query Match          12.6%   Score 200.5; DB 2; Length 281;  
Best Local Similarity 21.6%; Pred. No. 0.00036;  
Matches    72; Conservative      49; Mismatches 123; Indels     89; Gaps       10;
```

```
Qy  1 MEFGVAVTPKNATKVTPEDSLHVISQASLDCTVKSGESVVLSVTVGGAKLVIGTISQD 60  
Db  1 :||| ||:: :::::: :::: :::: :::: :::: :::: :::: :::: :::: :::  
Db  2 MOFFCTIVTKBEPPNVNLELDEGDIHFHTKAIIHPKSCKKGKVLIITA-----VISLMRED 54  
Qy  61 KFPQLSFDLVDFKEFELSHSGTSKANVHHFYGYKSNIEODDTSSDDDEDVPNAVPAAPTA 120  
Db  55 E:- - - - - :||| :::: -----MEEDDV--DOEESPREDIVEIPIG 78  
Qy  121 VTANG-----NAGAAVVADTKPKAKPAPKPARE-----EKPSDESDESDDD 163  
Db  79 ILEGKDIDDILNIYNG-QIWFELQAENGAYGVVALGSVITMGOGGCCDEDCDEH 137  
Qy  164 DESEDDDSKGMVDVEDSDDEESEDEEEETPKKPEPINKKRPNESVSKTPTVSGK 223  
Db  138 CLNHEDDE----EIDSDEFPGSDQEEDSDDELPLQIAPATKKGKITFISEVPKSKK 192  
Qy  224 KAKPAAAAASTPQTE-----KKKGCHTATPIPAKKGGKSPYNANOSP 266  
Db  193 EKTFEPKVPFKVEQVKOPTPOQQKAAAOOPEKANNKPAASAAPQNQSNAKPQP 252  
Qy  267 KSGGOS-SCGNNNKPFNSGKFQGSGNNKGNKSG 298  
Db  253 QOOQSPAKNNNRPNQNE-----NNKKQK 280
```

```
RESULT 14  
T09648  
nucleolin homolog num1 - alfalfa  
N:Alternate names: num1 protein  
C:Species: Medicago sativa (alfalfa)  
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000  
R:Accession: T09648  
R:Boegre, L.; Jonak, C.; Mink, M.; Meskienne, J.; Traas, J.; Ha, D.T.C.; Swoboda  
Plant Cell 8, 417-428, 1996  
A>Title: Developmental and cell cycle regulation of Alfalfa nucm1 a plant homo  
A:Reference number: 216796; MUID:96361876  
A:Accession: T09648  
A>Status: preliminary; translated from GR/EMBL/JDBJ  
A:Molecule type: mRNA  
A:Residues: 1-635 <BOB>  
A:Cross-references: EMBL:X88845; NID:g1279562; PID:NCAA61298.1; PID:g1279563  
C:Genetics:  
A:Gene: num1  
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
```

```
Query Match          12.4%   Score 196.5; DB 2; Length 635;  
Best Local Similarity 31.8%; Pred. No. 0.0013;  
Matches    56; Conservative      24; Mismatches 61; Indels     35; Gaps        6;
```

```
Qy  92 KSPNIEDDPFTSSDDDYVEAVPA-----PAPTAVTAGNACAANVVA---DTKPKAKPA 143  
Db  211 KDSSSEDSESSEDEDKKPAKASKNVSAPTKKAASSDESEDESEDEDAKPVSKPA 270  
Qy  144 EVYPAEEXPESEDESDD-----EDESEDDDSKGMVDVEDSDDD 186  
Db  271 AVAKSKKSDSDSDDEDDSSDDDKPKVASKEVSESESDDDDHK-MNIKDDSIDSD 329  
Qy  187 EEDESEDEEBETPKKPEPINKKRPNESVSKTPTVSGKKAKAAPAASTPQTOKTEKK 242  
Db  330 ESEESEDEPLTKPQK-----KIKOVMDV----AGKSGKKAPNTPATPNETSISK 376
```

Search completed: May 2, 2002, 22:07:31
Job time: 6679 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2002, 22:31:08 ; Search time 36.81 Seconds
(without alignments)
320.822 Million cell updates/sec

Title: US-09-645-337-8

Perfect score: 1589

Sequence: 1 MFWGVAVTPKNATKVTPEE.....KPGCGSNKCGSKGKGRA 305

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	311	19.6	412	1 FK84_SPOFR	Q26486 spodoptera
2	240	15.1	357	1 FK84_DROME	P54397 drosophila
3	215.5	13.6	694	1 NUCCL_CHICK	P15771 gallus gall
4	211.5	13.3	712	1 NUCCL_RAT	P13383 rattus norv
5	208	13.1	411	1 MP62_LYTP1	P91753 lytechinus
6	206.5	13.0	706	1 NUCCL_MOUSE	P09405 mus musculus
7	206	13.0	713	1 NUCCL_MESAU	P08199 mesocricetu
8	205	12.9	706	1 NUCCL_HUMAN	P19338 homo sapien
9	200.5	12.6	281	1 HS32_DICDI	P54658 dictyosteli
10	188	11.8	704	1 NP14_RAT	P41777 rattus norv
11	188	11.8	723	1 SSRP_DROME	Q05344 drosophila
12	183.5	11.5	294	1 NPM_CHICK	P16039 gallus gall
13	178.5	11.2	517	1 T2FA_HUMAN	P35269 homo sapien
14	177	11.1	644	1 NFM_RABIT	P54938 oryctolagus
15	177	11.1	650	1 NUCCL_XENLA	P20397 xenopus lae
16	175	11.0	524	1 T2FA_XENLA	Q04870 xenopus lae
17	173.5	10.9	361	1 FK84_SCHPO	Q74191 schizosacch
18	172	10.8	292	1 NPM_MOUSE	Q61937 mus musculus
19	172	10.8	678	1 GARP_PLAFL	P13816 plasmodium
20	170	10.7	1089	1 NMD2_YEAST	P38798 saccharomyc
21	169.5	10.7	543	1 TLPI_MOUSE	Q92273 mus musculus
22	169	10.6	292	1 NPM_RAT	P13084 rattus norv
23	168.5	10.6	299	1 NPM_XENLA	P07222 xenopus lae
24	168.5	10.6	699	1 NP14_HUMAN	Q14978 homo sapien
25	168	10.6	411	1 FK83_YEAST	P38911 saccharomyc
26	168	10.6	1002	1 YENA_DROME	P25992 drosophila
27	167.5	10.5	915	1 NFM_HUMAN	P01797 homo sapien
28	167	10.5	1275	1 TRP_DROME	P19334 drosophila
29	166	10.4	304	1 CEC1_CAEEL	P34618 caenorhabdi
30	166	10.4	542	1 TUL1_HUMAN	Q00294 homo sapien
31	164.5	10.4	472	1 YWIE_CAEEL	Q23525 caenorhabdi
32	164.5	10.4	589	1 RGPI_MOUSE	P46061 mus musculus
33	162.5	10.2	1109	1 TCF8_RAT	Q62947 rattus norv

Query Match

19.6% ; Score 311; DB 1; Length 412;

ALIGNMENTS

RESULT 1

FK84_SPOFR

ID FK84_SPOFR

STANDARD;

PRT; 412 AA.

AC Q26486;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-JUL-1999 (Rel. 38, Last annotation update)

DE 46 kDa FK506-binding nuclear protein (Peptidyl-prolyl cis-trans

DE isomerase) (PPIase) (EC 5.2.1.8).

OS Spodoptera frugiperda (Fall armyworm).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;

OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.

OX NCBI_taxid=7108;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95074110; PubMed=7527037;

RA Alnemri E.S., Fernandes-Alnemri T., Pomeroy K., Robertson N.M.,

RA Dudley K., Dubois G.C., Litwack G.,

RT *FKBP46, a novel Sf9 insect cell nuclear immunophilin that forms a

RT protein-kinase complex.*;

RL J. Biol. Chem. 269:30828-30834(1994).

CC - FUNCTION: BINDS TO, AND IS INHIBITED BY FK506 AND RAPAMYCIN. BINDS

CC - DOUBLE-STRANDED DNA IN VITRO. PPIASES ACCELERATE PROTEIN FOLDING.

CC - CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC

CC - PEPTIDE BONDS IN OLIGOPEPTIDES.

CC - SUBCELLULAR LOCATION: Nuclear.

CC - PTM: PHOSPHORYLATED BY A NUCLEAR KINASE IN THE PRESENCE OF MG2+

CC - AND ATP.

CC - SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@sib-sib.ch).

CC -----

CC EMBL; U15038; AAA58962.1; .

DR HSSP; P20071; 1FKJ.

DR InterPro; IPR001179; FKBP_PPIase.

DR Pfam; PF00254; FKBP; 1.

DR PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.

DR PROSITE; PS00454; FKBP_PPIASE_2; 1.

DR PROSITE; PS00509; FKBP_PPIASE_3; 1.

KW Isomerase; Rotamase; Nuclear protein; DNA-binding; Phosphorylation.

FT DOMAIN 91 112 ASP/GLU-RICH (HIGHLY ACIDIC).

FT DOMAIN 120 145 LYS-RICH (BASIC).

FT DOMAIN 152 216 ASP/GLU-RICH (HIGHLY ACIDIC).

FT DOMAIN 219 302 LYS-RICH (BASIC).

FT DOMAIN 324 412 PPIASE, FKBP-TYPE.

SQ SEQUENCE 412 AA; 45810 MW; F2A69159AEF4FE22 CRC64;

Db 102 TIENLLNSKAIKNSKKS-----EDDEDNESGEDEEDTDDDSOILBEYESFLENGEE 154

QY 168 EDDSEKGMVDVED--DSDDDEEDEDSEEE-EETPKPKPEPINKKRPNESVSKTPVSGKK 224
||||| ||||| :|||:||||| ||||| :|||:||||| ||||| :|||:|||||

Db 155 EDDDD-----DVDEDNESGEDEEDSDSDSEAEEOPK--AKVAKLSPGASAKK---SGKE 204

QY 225 AKPAAPASTPQTEKKKCGCHTATPHPAKKGKSPVNAQSPKS----- 268
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 205 ONGVAAKEAKOQOKKE-----KPEAKKEQPKAKPEAKQOPASKDPRTITGSKVIDQV 259
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 269 ---GGSSGG---NNKKPFNS---GKQF-----GGSNNKGSNKG 299
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 260 VNGEAKQKRVYVYIGRLQSNKTFDLSLLKGFKFKALGCGEVIKWDVG 312
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 3

NUCLCHICK STANDARD; PRT; 694 AA.

AC P15771;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nucleolin (Protein C23).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90206792; PubMed-2320420;

RA Marlor G., Nigg E.A.;

RT "cDNA sequences of chicken nucleolin/C23 and N038/B23, two major

RL Nucleolar proteins.";

RL Nucleic Acids Res. 18:1286-1286(1990).

RN [2]

RP DISCUSSION OF SEQUENCE.

RX MEDLINE-90304215; PubMed-2114180;

RA Marlor G., Nigg E.A.;

RT "Structure and developmental expression of chicken nucleolin and

RT N038: coordinate expression of two abundant non-ribosomal nucleolar

RL proteins.";

RL Biochim. Biophys. Acta 1049:126-133(1990).

RN [3]

RP SEQUENCE OF 407-694 FROM N.A.

RX MEDLINE-89119560; PubMed-2914325;

RA Borer R.A., Lehner C.F., Eppenberger H.M., Nigg E.A.;

RT "Major nucleolar proteins shuttle between nucleus and cytoplasm.";

RL Cell 56:379-390(1989).

CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING

CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR

CC CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN

CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A

CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.

CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.

CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X17199; CAA35060.1; -;

DR EMBL; M21791; AAA48983.1; -;

DR PIR; S08414; DNCNHL.

DR HSSP; P09651; IUP1.

DR InterPro; IPR000504; RRM.

DR Pfam; PF00076; rrm.4.

DR SMART; SM00360; RRM; 4.

DR PROSITE; PS50102; RRM; 4.

DR PROSITE; PS00030; RRM_RNP_1; 3.

KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;

RNA-binding.

FT DOMAIN 124 141 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 170 192 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 217 247 ASP/GLU-RICH (ACIDIC).

FT DOMAIN 281 357 RNA-BINDING (RRM) 1.

FT DOMAIN 371 445 RNA-BINDING (RRM) 2.

FT DOMAIN 461 535 RNA-BINDING (RRM) 3.

FT DOMAIN 553 628 RNA-BINDING (RRM) 4.

FT DOMAIN 650 682 ARG/GLY/PHE-RICH.

FT DOMAIN 55 90 5 X 7 AA TANDEM REPEATS OF X-T-P-X-K-K-X.

FT REPEAT 55 61 1.

FT REPEAT 62 68 2.

FT REPEAT 69 75 3.

FT REPEAT 76 82 4.

FT REPEAT 84 90 5.

FT MOD_RES 116 116 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 136 136 PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 171 171 PHOSPHORYLATION (BY SIMILARITY).

FT CONFLICT 419 419 A -> R (IN REF. 3).

FT CONFLICT 520 520 N -> T (IN REF. 3).

SQ SEQUENCE 694 AA; 75640 MW; 7996C504BE9459A1 CRC64;

Query Match 13.6%; Score 215.5; DB 1; Length 694;

Best Local Similarity 28.1%; Pred. No. 5.2e-05;

Matches 85; Conservative 32; Mismatches 102; Indels 83; Gaps 12;

QY 6 VAVTKNATK-----VTPEEDSLVHSQASLDCTVSGESVWL-----SVT----- 46
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 4 LAKTPKNQMKQKMAPPKPKKEEESSEESDLEESSEGVMPKPKQKAAVTPAKKAAT 63

QY 47 -----VGGAKLVIGTSLQDKFPQISFDLVDFKFEFELSHSGTKANVHFIGYKSP 94
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 64 PAKKAATPAKKAATPAKKAAT-----PAKKAVAPSPKKAAVGVCKGAKNGKNKKEESE 117

QY 95 NIEQDDFTSSDDDEVPAPAPATAVTANGNAGAAVVKADTKPKAKPAEVKPAKEKPS 154
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 118 EDE 171

QY 155 DEDESDDE 187
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 172 EEEEEDE 231

QY 188 EEDSEDEEEETPKPEPINK---KRPNESVSKTPVSGKKAKPAAPAS-----TP 235
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db 232 EDE 290

QY 236 OK 237

Db 291 TK 292

RESULT 4

NUCL_RAT STANDARD; PRT; 712 AA.

AC P13383;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Nucleolin (Protein C23).

GN NCL OR NUC.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-90269607; PubMed-2347493;

RA Bourbon H.-M., Amalric F.;

RT "Nucleolin gene organization in rodents: highly conserved sequences

RT within three of the 13 introns.";

QY 145 VKPAEKPESEDE-----ESDDESEDEDDSEKGMVDDEDDSD 185
DB 241 NGAHIDDEDEDEDEYKVGDEEEERATSGEEEEEDEEEDEEMALG-DDDDEDD 299
QY 186 DEEE-----DSDEEEETPKPKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKT 238
DB 300 DEEDDEDEGMDDEDEEEESSVPKPAK-----AKGVNG-TAKPKGTPKSQANKG 351
QY 239 EKKK 242
DB 352 MKEK 355
RESULT 6
ID NUCLE_MOUSE STANDARD; PRT; 706 AA.
AC P09405; Q61991;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL OR NUC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=88316930; PubMed=3137346;
RA Bourbon H., Lapeyre B., Amalric F.;
RT "Structure of the mouse nucleolin gene. The complete sequence reveals
that each RNA binding domain is encoded by two independent exons.";
RL J. Mol. Biol. 200:627-638(1988).
RN [2]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=89121496; PubMed=2906027;
RA Pasternack M.S., Bleier K.J., McInerney T.N.;
RT "Granzyme A binding to target cell proteins. Granzyme A binds to and
cleaves nucleolin in vitro.";
RL J. Biol. Chem. 266:14703-14708(1991).
CC -!- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CHROMATIN AND PRERIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
ROLE IN PRE-RRNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -!- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@lsb-sib.ch).
CC -----
DB EMBL; X07699; CAA30538.1; -;
DB EMBL; M22089; AAA39841.1; -;
DB PIR; A29588; DNMS.
DB HSSP; P09651; 1HA1.
DB MGD; MGI:97286; Ncl.
DB InterPro; IPR000504; RRM.
DB Pfam; PF00076; rrm; 4.
DB SMART; SM00360; RRM; 4.

DR PROSITE: PS50102; RRM; 4.
DR PROSITE: PS00030; RRM_RNP_1; 3.
KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
FT RNA-binding.
FT INIT_MET 0
FT DOMAIN 142 169 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 189 214 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 240 272 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 308 384 RNA-BINDING (RRM) 1.
FT DOMAIN 394 467 RNA-BINDING (RRM) 2.
FT DOMAIN 486 560 RNA-BINDING (RRM) 3.
FT DOMAIN 568 643 RNA-BINDING (RRM) 4.
FT DOMAIN 645 696 ARG/GLY/PHE-RICH.
FT MOD_RES 144 144 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 156 156 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 188 188 PHOSPHORYLATION (BY SIMILARITY).
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-X-X.
FT REPEAT 57 64 1.
FT REPEAT 74 81 2.
FT REPEAT 82 89 3.
FT REPEAT 90 97 4.
FT REPEAT 98 103 5 (INCOMPLETE).
FT REPEAT 104 111 6.
FT REPEAT 119 126 7.
FT REPEAT 127 134 8.
SQ SEQUENCE 706 AA; 76592 MW; 89505ER39C89F832 CRC64;
Query Match 13.0%; Score 206.5; DB 1; Length 706;
Best Local Similarity 33.0%; Pred. No. 0.00016;
Matches 65; Conservative 15; Mismatches 54; Indels 63; Gaps 9;
QY 97 EDDFTSSDDDEVPAPAPAPATVANGAAGVAVKADTKPKAKPAKVPKPAEKPESEDE 156
DB 150 EEDEDDSEDEDEDEDEFEPP-----IVKG-VKP-AKAPAPASEDEDEDE 195
QY 157 -EDESDEDESEDDSEKGMV-----DEDDSD-D 185
DB 196 DEDEDDDEDEDEDESEEVMEITAKGKTPAKVVPKAKSAVEEDDEDEDEDEDEDE 255
QY 186 DEEDSEDE-----EEEEETPKPKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEKK 242
DB 256 DEEDDEDEDEDEDEDEDEEPVKAAPGKRKK-----EMTKQEAPEAKKQKVE--- 300
QY 243 GGTATPHPAKKGKSP 259
DB 301 GSEPTTFNLFIGNLP 317
RESULT 7
ID NUCLE_MOUSE STANDARD; PRT; 713 AA.
AC P08199;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nucleolin (Protein C23).
GN NCL.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE OF 35-713 FROM N.A., AND SEQUENCE OF 1-34.
RX MEDLINE=87175501; PubMed=3470736;
RA Lapeyre B., Bourbon H., Amalric F.;
RT "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
an unusual protein structure revealed by the nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
RN [2]
RP FUNCTION.

RX MEDLINE-88312631; PubMed-3409881;
RA Erard M.S., Helenguer P., Caizergues-Ferrer M., Pantaloni A.,
RA Analric F.;
RT "A major nucleolar protein, nucleolin, induces chromatin
RT decondensation by binding to histone H1.";
RL Eur. J. Biochem. 175:525-530(1988).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PRE-RIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M15825; AAA36966.1; -
CC PIR: A27441; A27441.
CC HSP; P09651; 1HA1.
CC InterPro: IPR000504; RRM.
CC Pfam: PF00076; rrm; 4.
CC SMART: SM00360; RRM; 4.
CC PROSITE: PS50102; RRM; 4.
CC PROSITE: PS00030; RRM_RNP_1; 3.
CC Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
CC RNA-binding.
CC INIT_MET 0 0
CC DOMAIN 141 169 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 188 213 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 239 271 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 307 383 RNA-BINDING (RRM) 1.
CC FT DOMAIN 393 466 RNA-BINDING (RRM) 2.
CC FT DOMAIN 485 559 RNA-BINDING (RRM) 3.
CC FT DOMAIN 571 646 RNA-BINDING (RRM) 4.
CC FT DOMAIN 648 701 ARG/GLY/PHE-RICH.
CC FT DOMAIN 56 133 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
CC FT REPEAT 56 63 X-X.
CC FT REPEAT 73 80 1.
CC FT REPEAT 81 88 2.
CC FT REPEAT 89 96 3.
CC FT REPEAT 97 102 4.
CC FT REPEAT 103 110 5 (INCOMPLETE).
CC FT REPEAT 118 125 6.
CC FT REPEAT 126 133 7.
CC FT MOD_RES 143 143 X-T-P-X-K-K-X-X MOTIF.
CC FT MOD_RES 156 156 PHOSPHORYLATION.
CC FT MOD_RES 187 187 PHOSPHORYLATION.
CC FT MOD_RES 655 655 METHYLATION (DI-).
CC FT MOD_RES 659 659 METHYLATION (DI-).
CC FT MOD_RES 665 665 METHYLATION (DI-).
CC FT MOD_RES 669 669 METHYLATION (DI-).
CC FT MOD_RES 673 673 METHYLATION (DI-).
CC FT MOD_RES 679 679 METHYLATION (DI-).
CC FT MOD_RES 681 681 METHYLATION (DI-).
CC FT MOD_RES 687 687 METHYLATION (DI-).
CC FT MOD_RES 691 691 METHYLATION (DI-).
CC FT MOD_RES 694 694 METHYLATION (DI-).
CC SEQUENCE 713 AA; 76997 MW; 79DDCF724CED7DB4 CRC64;

Query Match 13.0%; Score 206; DB 1; Length 713;
Best Local Similarity 23.2%; Pred. No. 0.00017;
Matches 66; Conservative 31; Mismatches 81; Indels 106; Gaps 8;
QY 97 EQDDFTSSDDHEDVPE-----AVPAPA----- 117

DB 34 EEDSSGSEVVIPQKKKKKATATPAKKVVVSQTKKVAVTPPAKAAVTPGKKAATPAK 93
QY 118 -----PTAVTANGAAGAAVVDTKPKAKPAEVKPAE-----EKPESEDESDDEDE 165
DB 94 KAVTPAKAVATPGKKGATOKALVATPGKKGAVTPAKGAKNGKNAKKEDDDDDDE 153
QY 166 SEEDDSEKGMVDREDD-----SDDDEDESEDEFEFEETPKKPEP 205
DB 154 DSDSEDEE---DEEDEFEPVVGKQGVAAAAPASEDEDEDEDEE-----EE 205
QY 206 INKRRPNESVSTPVSCKKAKPAAPAST-----PQKTE 239
DB 206 EDDSEEEAEITPAKKGKAPKAVPVVPAKNAEEDDDDEDEDEDEDEDEDEDEE 265
QY 240 KKGKGTATPHPAKKGKSPVNAQSPKSGQSSGGNNKKPN 283
DB 266 EEEEPVPAPGKR-KKMTKQKEVPEAKKQKVEGSESTTPN 308

RESULT 8
NUCL_HUMAN
ID NUCL_HUMAN STANDARD; PRT; 706 AA.
AC P19338;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nucleolin (protein C23).
GN NCL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89290043; PubMed-2737305;
RA Srivastava M., Fleming P.J., Pollard H.B., Burns A.L.;
RT *Cloning and sequencing of the human nucleolin cDNA.*;
RL FEBS Lett. 250:99-105(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-90368666; PubMed-2394707;
RA Srivastava M., McBride O.W., Fleming P.J., Pollard H.B., Burns A.L.;
RT *Genomic organization and chromosomal localization of the human
RT nucleolin gene.*;
RL J. Biol. Chem. 265:14922-14931(1990).
CC -1- FUNCTION: NUCLEOLIN IS THE MAJOR NUCLEOLAR PROTEIN OF GROWING
CC EUKARYOTIC CELLS. IT IS FOUND ASSOCIATED WITH INTRANUCLEOLAR
CC CHROMATIN AND PRE-RIBOSOMAL PARTICLES. IT INDUCES CHROMATIN
CC DECONDENSATION BY BINDING TO HISTONE H1. IT IS THOUGHT TO PLAY A
CC ROLE IN PRE-RNA TRANSCRIPTION AND RIBOSOME ASSEMBLY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: CONTAINS 4 RNA RECOGNITION MOTIFS (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60858; AAA59954.1; -
CC PIR: A35804; A35804.
CC PIR: S04631; S04631.
CC HSP; P09651; 1HA1.
CC Aarhus/Chent-2DPAGE; 1210; NEPHGE.
CC MIM; 164035; -
CC InterPro: IPR000504; RRM.
CC Pfam: PF00076; rrm; 4.
CC SMART: SM00360; RRM; 4.
CC PROSITE: PS50102; RRM; 4.
CC PROSITE: PS00030; RRM_RNP_1; 3.

```
KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding.
FT INIT_MET 0
FT DOMAIN 142 170 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 184 208 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 233 270 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 306 382 RNA-BINDING (RRM) 1.
FT DOMAIN 332 465 RNA-BINDING (RRM) 2.
FT DOMAIN 485 559 RNA-BINDING (RRM) 3.
FT DOMAIN 571 643 RNA-BINDING (RRM) 4.
FT DOMAIN 645 694 ARG/GLY/PHE-RICH.
FT DOMAIN 57 134 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
FT REPEAT 57 64 X-X.
FT REPEAT 74 81 1.
FT REPEAT 82 89 2.
FT REPEAT 90 97 3.
FT REPEAT 98 103 4.
FT REPEAT 104 111 5 (INCOMPLETE).
FT REPEAT 119 126 6.
FT REPEAT 127 134 7.
FT MOD_RES 144 144 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 152 152 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 183 183 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 706 AA; 76213 MW; 85A2F2CA22EA03DB CRC64;

Query Match 12.9%; Score 205; DB 1; Length 706;
Best Local Similarity 27.2%; Pred. No. 0.00019;
Matches 74; Conservative 29; Mismatches 95; Indels 74; Gaps 9;

QY 6 VAVTP-KNATKVTPEEDSLVHISQASLDCTVKSQSVLVTVG-----AKLVIGTILS 58
DB 63 VVVSPTKKVAVATPAKKAATPGKKAATPAKKTTPAKAVITPGKKGATPGKALVATPG 122
QY 59 QDKFPQISDLVDFKEFELSHSGTKANVHFIGYKSPNIEDDF--TSSDDRDVPEAVPAP 116
DB 123 K-----KGAAIPAKAGKGNKAKEDSDEEDDDDEEDDEDEDEDEDEDEDE 169
QY 117 APTAVTANGNAGVAVKADTKPKAPKPAEVPKPEPSEDESDDESE-----168
DB 170 EPAAMKA-----PASEDEDEDEDEDEDEDEDEDEDEDEDEDESEBAMETTPA 215
QY 169 -----DDSEKGMVDVDESDSDDEEDSDESEDESEDESEDESEDESEDESE 209
DB 216 KGKAAKVPVPAKNAVDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 271
QY 210 RPNEYSKTPVSGKK--AKPAAAPASTPKTE 239
DB 272 -----VKEAPGKRRKEMAKOKAAPEAKKQKVE 298

RESULT 9
ID HS32_DICDI STANDARD; PRT; 281 AA.
AC P34658;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 32 kDa heat shock protein (4-1 protein).
GN HSPC OR HSP32.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96001265; PubMed=7557471;
RA de Maria A.C., Gomes S.L., Juliania M.H., Mazzarella R., Klein C.;
RT *Cloning of a cDNA encoding a novel heat-shock protein from
RT Dictyostelium discoideum.*
RL Gene 163:163-164(1995).
CC -I- DEVELOPMENTAL STAGE: PRESENT AT HIGH LEVELS IN GROWING CELLS BUT
DECREASES DRAMATICALLY DURING THE EARLY HOURS OF DEVELOPMENT.
```

```
CC -I- INDUCTION: BY HEAT SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40211; AAC47710.1; -.
DR EMBL; L39778; AAA99510.1; -.
DR Dictydb; DD01076; hspC.
KW Heat shock. 251 257 POLY-GLN.
FT DOMAIN 262 265 POLY-ASN.
SQ SEQUENCE 281 AA; 31481 MW; 2C64F57C847D7CA4 CRC64;

Query Match 12.6%; Score 200.5; DB 1; Length 281;
Best Local Similarity 21.6%; Pred. No. 0.00014;
Matches 72; Conservative 49; Mismatches 123; Indels 89; Gaps 10;

QY 1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSQSVLVTVGAKLVIGTILSQD 60
DB 2 MOPFGTIVTKPEPVNLELDEGDIFHLTKAIHPKSGKGKLYLTA-----VISLMEED 54
QY 61 KFPQISDLVDFKEFELSHSGTKANVHFIGYKSPNIEDDTSSDDEVDVPEAVPAPATA 120
DB 55 E-----MEEDV--DDEESPREDIVPEIG 78
QY 121 VTANG-----NAGAAVVKADTKPKAPKPAEVPKPAE-----EKPESEDESDDE 163
DB 79 ILEAGKIDQIDLNLHYNFG-QIVREFLOAENGAGYVVALSGSVITWMEGGCDDCDDH 137
QY 164 DESEDDDDSEKGMVDVDESDSDDEEDSDEDEEBETPKKPEPKPKRPNESKTPVSGK 223
DB 138 CINHEDDE-----EIDSDDEFGSDDEEDSDEDEIPQLIATPKKGITETISEVPESKK 192
QY 224 KAKPAAAPASTPKTE-----KKGGHTATPHPAKKGKSPVNAOSP 266
DB 193 EKTPEKPKVPKPEKQVKOPTQPKQKAAAOQPEKANNKPAASPAKPNQNSKNAPQP 252
QY 267 KSGGGS-SGNNKPKFPNSGKQFGSGSNK 298
DB 253 QQQQSPAKNNNKRPNQNE-----NNKKQK 280

RESULT 10
NP14_RAT
ID NP14_RAT STANDARD; PRT; 704 AA.
AC P41777;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa
DE nucleolar phosphoprotein) (Nop140) (Nucleolar and coiled-body
DE phosphoprotein 1).
GN NOLC1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.
RC TISSUE=Liver;
RX MEDLINE=92323542; PubMed=1623516;
RA Meier U.T., Blobel G.;
RT *Nop140 shuttles on tracks between nucleolus and cytoplasm.*;
RL Cell 70:127-138(1992).
RN [2]
RP INTERACTION WITH NOP5 AND FIBRILLARIN.
RX MEDLINE=20143579; PubMed=10679015;
```



```

Db 135 KIGN-----ASTKRPSGGGAKTPOKPKLSDEDDDEDDDEDD-----DEDD 179
Qy 183 SDDDEEDSEDEEETP-KKP--EPINKRPNESVSKTPVSGKAKPAAPASTPQKTE 239
Db 180 LDO-----DEEIKTPMKKPAEPAGK-----NMQKAKQNGKSKSPSPASKT--KTP 225
Qy 240 KKGCGHTATPHAPK 253
Db 226 DSKDKSLTPKTPK 239

RESULT 13
T2FA_HUMAN
ID T2FA_HUMAN STANDARD; PRT; 517 AA.
AC P35269;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription initiation factor IIF, alpha subunit (TFIIF-alpha)
DE (transcription initiation factor RAP74).
GN GTF2F1 OR RAP74.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-92131135; PubMed-1734284;
RA Aso T., Vasavada H.A., Kawauchi T., Germino F.J., Ganguly S.,
RA Kitajima S., Weissman S.M., Yasukochi Y.;
RT "Characterization of cDNA for the large subunit of the transcription
RT initiation factor TFIIF.";
RL Nature 355:461-464(1992).
[2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-92131136; PubMed-1734284;
RA Finkelstein A., Kostrub C.F., Li J., Chavez D.P., Wang B.O.,
RA Fang S.M., Greenblatt J., Burton Z.F.;
RT "A cDNA encoding RAP74, a general initiation factor for transcription
RT by RNA polymerase II.";
RL Nature 355:464-467(1992).
CC -1- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT
CC BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE
CC INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES
CC TRANSCRIPTION ELONGATION.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X64037; CAA45408.1; -
DR EMBL; X64002; CAA45404.1; -
DR PIR; S20248; S20248.
DR TRANSFAC; T02168; -
DR MM; 189968; -
DR Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation.
FT CONFLICT 231 V -> I (IN REF. 2).
FT CONFLICT 361 F -> L (IN REF. 2).
FT SEQUENCE 517 AA; 58274 MW; F0D2BE44D2F3820F CRC64;
SQ
Query Match 11.2%; Score 178.5; DB 1; Length 517;
Best Local Similarity 26.8%; Pred. No. 0.0036;
Matches 74; Conservative 34; Mismatches 83; Indels 85; Gaps 13;

```

```

Qy 72 DKFELSHGSTRANVHFIYKSPNIQDDFTSSDDHVPVPAVPAVPAVANGAGA 131
Db 193 EKEKRRRKASSELRII-----DLEDDLEMSDASD-----ASGEGGRV 231
Qy 132 VKADTK-PRAKPAEVKPAEKPEDEE--DESDDEI-----ESFDDDSK- 173
Db 232 PRAKKAPLAKGRKK--KKKGSDDPEAFSDGDFKQGVYMSIDSGSSSOERPEKA 289
Qy 174 -----KGMVDDEDDSDDEEE--DSEDEEEETPKKPEPINKKRPNESVSKTPVS- 221
Db 290 KAPQOEKPGKGVDEQSDSESESESEKPEDEKKEEKKAPTQPKKKRKKDSKESDSE 349
Qy 222 -----GKAKPA--AAPASTPQTEKKKGHTATPH-----PAKKGCKS- 258
Db 350 ESDIDSEASSAFFMAKKKTPPKRERKPSGSGSRGSRGPTPSAEGGSTSTLRAAAKLE 409
Qy 259 -----PVNANQSPKSGSGSGGNNKKPFNSCK 286
Db 410 QGKRSEMPAAKRLRLDTGQSLSGKSTQPP-PSGK 444

RESULT 14
NFM_RABIT
ID NFM_RABIT STANDARD; PRT; 644 AA.
AC P54938;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet M protein (160 kDa neurofilament protein)
DE (Neurofilament medium polypeptide) (NF-M) (fragment).
GN NEF3 OR NFM OR NFM.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE-97055255; PubMed-8899542;
RA Vitadello M., Vettore S., Lamar E., Chien K.R., Gorza L.;
RT "Neurofilament M mRNA is expressed in conduction system myocytes of
RT the developing and adult rabbit heart.";
RL J. Mol. Cell. Cardiol. 28:1833-1844(1996).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFM IS
CC PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS
CC THOUGHT THAT PHOSPHORYLATION OF NFM RESULTS IN THE FORMATION OF
CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
CC OF AXONAL CALIBER.
CC -1- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
CC OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
CC LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
CC COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z47378; CAA87454.1; -
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone; Phosphorylation;
FT NON_TER 1 1
FT DOMAIN <1 197 ROD.

```


Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1589	100.0		305	10	Q49209	Q49209 arabidopsis
2	1578.9	99.3		306	10	Q9FNJ6	Q9fnj6 arabidopsis
3	594	37.4		245	10	Q9FVE6	Q9fve6 arabidopsis
4	594	37.4		257	10	Q22238	Q22238 arabidopsis
5	574	36.1		297	10	Q9MAT5	Q9mat5 oryza sativ
6	543.5	34.2		307	10	Q24591	Q24591 zea mays (m
7	541	34.0		303	10	Q9MAU5	Q9mau5 zea mays (m
8	532	33.5		300	10	Q9MAU4	Q9mau4 zea mays (m
9	453.5	28.5		294	10	Q9MAT4	Q9mat4 arabidopsis
10	450.5	28.4		296	10	Q9L2R5	Q9l2r5 arabidopsis
11	364	22.9		286	10	Q94F81	Q94f81 zea mays (m
12	262.5	16.5		203	10	Q9MAT3	Q9mat3 arabidopsis
13	258	16.2		312	5	Q26711	Q26711 trypanosoma
14	230	14.5		798	5	Q45181	Q45181 caenorhabdi
15	210.5	13.2		346	5	Q9NLA3	Q9nla3 asterina pe
16	207.5	13.1		611	10	Q41042	Q41042 pisum sativ

QY	1 MEFWGVATPKNATKVTPEEDSLVHISOASLDC ^T VKSGSVVL SVTVGAKLVIGTLSQD 60
Db	1 MEFWGVATPKNATKVTPEEDSLVHISOASLDC ^T VKSGSVVL SVTVGAKLVIGTLSQD 60

```
QY 61 KFPQISFDLVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 120
Db 61 KFPQISFDLVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 120
QY 121 VTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 180
Db 121 VTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 180
QY 181 DSDDDDEEDSEDEEETPKKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
Db 181 DSDDDDEEDSEDEEETPKKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
QY 241 KKGHTATPHPAKGGKSPVNAQSPKSGGSGGNNKKPFNSGKQFGGNNKGSNKGK 300
Db 241 KKGHTATPHPAKGGKSPVNAQSPKSGGSGGNNKKPFNSGKQFGGNNKGSNKGK 300
QY 301 GKGRA 305
Db 301 GKGRA 305
RESULT 2
Q9FNJ6 PRELIMINARY; PRT; 306 AA.
AC Q9FNJ6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HISTONE DEACETYLASE-LIKE PROTEIN.
GN A75G22650 OR MJ22.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones."
RL DNA Res. 4:291-300(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB006699; BAB1671.1; -
DR EMBL; AY059893; AAL24375.1; -
SQ SEQUENCE 306 AA; 32348 MW; A19274D43BDD142C CRC64;
```

Query Match 99.3%; Score 1578.5; DB 10; Length 306;
Best Local Similarity 99.7%; Pred. No. 9e-98; Mismatches 0; Indels 1; Gaps 1;
Matches 305; Conservative 0

```
QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCITVKSGESVLSVTVGGAKLVIGTISQD 60
Db 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCITVKSGESVLSVTVGGAKLVIGTISQD 60
QY 61 KFPQISFDLVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 120
Db 61 KFPQISFDLVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 120
QY 121 VTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 180
Db 121 VTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 180
```

```
Db 121 VTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 180
QY 181 DSDDDDEEDSEDEEETPKKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 239
Db 181 DSDDDDEEDSEDEEETPKKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 240
QY 240 KKGHTATPHPAKGGKSPVNAQSPKSGGSGGNNKKPFNSGKQFGGNNKGSNKGK 299
Db 241 KKGHTATPHPAKGGKSPVNAQSPKSGGSGGNNKKPFNSGKQFGGNNKGSNKGK 300
QY 300 GKGRA 305
Db 301 GKGRA 306
RESULT 3
Q9FVE6 PRELIMINARY; PRT; 245 AA.
AC Q9FVE6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN HD2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wu K., Tian L., Malik K., Brown D., Miki B.;
RT "Functional analysis of HD2 histone deacetylase homologs in
RT Arabidopsis thaliana."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195545; AAG28472.1; -
DR InterPro; IPR000822; Znf-C2H2.
DR SMART; SM00355; Znf-C2H2.1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2.1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 245 AA; 26372 MW; E8E4F7E0A0D24F28 CRC64;
```

Query Match 37.4%; Score 594; DB 10; Length 245;
Best Local Similarity 50.7%; Pred. No. 2.4e-32;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

```
QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCITVKSGESVLSVTVGGAKLVIGTISQ 59
Db 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCITVKSGESVLSVTVGGAKLVIGTISQ 60
QY 60 DKFPQISFDLVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 119
Db 61 ENIPQICDILVDFKFEFELSHSGTKANVHFICGYKSPNIEQDDFTSSDDEDDVPEAVPAPAPTA 115
QY 120 AVTANGAGAAVVKADTKPKAKPAEVKPAEKPESDEDESDDESDDESDDESDDESDDESDDESD 179
Db 116 -----GNAKAV-----AKPKAPAEVKPAVD-----DEKDESUD-----GM--D 149
QY 180 EDDSDDEEDSEDEEETPKKPEINKKRPNESVSKTPVSGKKAKPAAAPASTPQKTEK 239
Db 150 EDDSDG---EDS---EEETPKKPAK---SKKRANETTPKAPVSAKAKAVAV-----TPQKTD 199
QY 240 KKGHTATPHPAKGGKSPVNAQSPKSGGSGGNNKKPFNSGKQFGGNNKGSNKGK 297
Db 200 EK-----KKGKA---ANOSPKSASQVSCG-SCKKTFNSGNAL--RSHNRAKH 241
RESULT 4
ID O22238 PRELIMINARY; PRT; 257 AA.
```

```

AC 02238:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE.
GN T32N15.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA de la Bastide M.R., Parnell L.D., Kaplan N., Gnoj L., Hameed A.,
RA Schütz K., Hasegawa A., Gottesman T., Shohdy N., Granat S., Jensen K.,
RA Johnson A.F., Lodhi M., Dedhia N., Martienssen R., McCombie W.R.;
RT "A. thaliana BAC T32N15 from chromosome V.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002534; AAB70032.1; -.
DR InterPro: IPR000822; Znf-C2H2.
DR Pfam: PF00096; zf-C2H2; 1.
DR SMART: SM00355; Znf-C2H2; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
DR DNA-binding; Metal-binding; Zinc-finger.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 257 AA; 27698 MW; C4404387686889D5 CRC64;

Query Match 37.4%; Score 594; DB 10; Length 257;
Best Local Similarity 50.7%; Pred. No. 2.5e-32;
Matches 151; Conservative 26; Mismatches 63; Indels 58; Gaps 15;

QY 1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSSESVLSTVGGAKLVIGTLSD 59
Db 1 MEFWGLEVRKPGVTPTEEGILHVSQASLGCKNKGFEVPLHVKNQNLVGLTST 60
QY 60 DKFPQISFDLVDFKEFELSHSGTKANVHFVIGYKSPNIEQDDFTSSDDEVPAPAPAT 119
Db 61 ENTPQIFDLVDFKEFELSHSGTKANVHFVIGYKSPNIEQGY-SEEEEEEVPA- 115
QY 120 AVTANGNAGAAVYKADTKPKAKPAEVKPAEKPESEDEDEDEDEDEDEDEDEDE 179
Db 116 -----GNAKAV-----AKPKAKPAEVKPAVD-----DEEDSD-----GW-D 149
QY 180 EDDSDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 239
Db 150 EDDSDG---EDS---EEETPPKPPAS-SKKRANETTPKAPVSAKAKAVAV---TPQKTD 199
QY 240 KKGKGGTATPHPAKKGKSPVNAQSPKSGGSGGNNKPPNSGKQFGGNNKGSN 297
Db 200 EK-----KGGKA---ANQSPKSAQVSCG-SCKKTFNSGNAL-ESHNAKH 241

RESULT 5
Q9M4T5 PRELIMINARY; PRT: 297 AA.
ID Q9M4T5
AC Q9M4T5
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE HD2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Dangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT "Molecular characterization of type-2 histone deacetylases in higher
RT plants.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL: AF255711; AAF70196.1; -.
DR InterPro: IPR000822; Znf-C2H2.
DR SMART: SM00355; Znf-C2H2; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 297 AA; 32502 MW; EEB3AAC0D2C41C63 CRC64;

Query Match 36.1%; Score 574; DB 10; Length 297;
Best Local Similarity 46.0%; Pred. No. 6.2e-31;
Matches 137; Conservative 38; Mismatches 93; Indels 30; Gaps 11;

QY 1 MEFWGVAVTPKNATKVTPEEDSLVHISQASLDCTVKSSESVLSTVGGAKLVIGTLSD 60
Db 1 MEFWGLEVRKPGVTPTEEGILHVSQASLGCKNKGFEVPLHVKNQNLVGLTST 60
QY 61 KFPQISFDLVDFKEFELSHSGTKANVHFVIGYK-SPNIEQD--DFTS---SDDEDVPEAVP 114
Db 61 KFPQIFDLVDFKEFELSHSGTKANVHFVIGYKSPNIEQD--DFTS---SDDEDVPEAVP 120
QY 115 APAPTAVTANGNAGAAVYKADTKPKAKPAEVKPAEKPESEDEDEDEDEDEDEDEDE 172
Db 121 AP-----RANGK-----VEGKENEQKQKGTSSASKAAVNDODDDSDAEDSDQED 170
QY 173 EKGMDVDEDDSDDEDEDEDEDE-EEETPPKPEPINKRPNESVSKTPVSGKAKAPAAAP 231
Db 171 LSPEDDDDDSDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 226
QY 232 ASTP--OKTEKKKGHTATPHPAKKGKSPVN---ANQSPKSGGSGGNNKPPNS 284
Db 227 ---TPSQKGTGDKKGVHATPHPAKQAQSKTPVNDKSKESPKSGGSGISCKSCSKTFNS 282

RESULT 6
O24591 PRELIMINARY; PRT: 307 AA.
ID O24591
AC O24591
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NUCLEOLAR HISTONE DEACETYLASE HD2-P39.
GN HD2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CUZCO 251;
RX MEDLINE=97349336; PubMed=9204905;
RA Lusser A., Brosch G., Loidl P., Haas H., Loidl P.;
RT "Identification of maize histone deacetylase HD2 as an acidic
RT nucleolar phosphoprotein.";
RL Science 277:88-91(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CUZCO;
RA Brosch G., Lusser A., Loidl P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U82815; AAB63262.1; -.
DR EMBL: AF026917; AAC61674.1; -.
DR PROSITE: IPR000822; Znf-C2H2.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger.
SQ SEQUENCE 307 AA; 33238 MW; 9CFF6B036D274FC1 CRC64;

Query Match 34.2%; Score 543.5; DB 10; Length 307;
Best Local Similarity 44.8%; Pred. No. 6.9e-29;
Matches 130; Conservative 45; Mismatches 96; Indels 19; Gaps 11;

```

```

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCTVKSQSVLVTGGAKLVIGTILSD 60
Db 1 MEFWGLEVKPGSTVKGPGYGVFLHLSQAALGESKKS-DNALMYVKIDQKLAIGTILSD 59

QY 61 KFOISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDDFTSSDDDEVDPEAVPAPAPTA 120
Db 60 KNPHIQDFLDFKFEFELSHSGTKSVFTTCYKVPQPEDEMDLSDDEDEELNVP---V 116

QY 121 VTANGNAGAAVVKADTKPKAKPAEYKPAEKPESEDESEDESEED--DDSEKGMV 178
Db 117 VKENGKADKKQKQSKAVAAAPSKSSPDKSKDDDDSDDETDSDDEDETDSDDEGLSS 176

QY 179 DEEDSDDEDEEDSDDEEE--TPKKPEPINKKRPNE-SYKSTPVSCKKAKPAAPASTP 235
Db 177 EGGDDSDDEEDTSDDEEDTTPKPE-VGKKRPAESSVLKTPLSDKKAK VATPSS-- 232

QY 236 QKTEKKKGG--HTATPHAKKGGKSPVNAQSPKS--GGSSGNNKKP 281
Db 233 QKTGKKGAHVATPHPAK--GKTIVNNDKSVKSPKSAKSGSVCPCKP 280

RESULT 7
QY4U5 PRELIMINARY; PRT; 303 AA.
ID QY4U5;
AC QY4U5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE HISTONE DEACETYLASE 2 ISOFORM B.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OX Panicoideae; Andropogoneae; Zea.
RN NCBI_TaxID=4577;
RP SEQUENCE FROM N.A.
RA Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT "Molecular characterization of type-2 histone deacetylases in higher
plants.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254073; AAF68624.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2.1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger;
SQ SEQUENCE 303 AA; 70854627101BB67C CRC64;

```

```

Query Match 34.0%; Score 541; DB 10; Length 303;
Best Local Similarity 44.1%; Pred. No. 9.9e-29;
Matches 135; Conservative 36; Mismatches 95; Indels 40; Gaps 12;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCTVKSQSVLVTGGAKLVIGTILSD 60
Db 1 MEFWGLEVKPGSTVKGPGYGVFLHLSQAALGESKKS-DNALMYVKIDQKLAIGTILSD 59

QY 61 KFOISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDDFTSSDDDEVDPEAVPAPAPTA 120
Db 60 KYPQIDLVFNKFEFELSHSGTKSVFTTCYKVPQPEDEMDLSDDEDEELNVP---V 116

QY 121 VTANGNAGAAVVKADTKPKAKPAE-----VKPAEKPESEDESEDESDS 166
Db 117 IKENG-----KADGKEQKQKAVAAATASKSLGLEKSKSDSDSDSDSDSDSD 169

QY 167 EEDSDS-KGMDVDEDDSDDEEDSDDEEE--TPKKPEPINKKRPNE-SYKSTPVSCK 223
Db 170 DSDSDSEGLSPDGGDDSDDEEDTSDDEEDTTPKPE-ACKKGAENALKTPLSDK 228

QY 224 KAKPAAAPASTPQTEKKKG--HTATPHAKKGGKSPVNAQSPKS-----ANOSPKSGSGSGNNN 278
Db 229 KAKVATPPA---QKTGCKKGATHVATPHPAK--GKTIVNNDKSVKSPKSAKSG-SVPCCKSC 282

```

```

QY 279 KKPFS 284
Db 283 SKTFS 288

RESULT 8
QY4U4 PRELIMINARY; PRT; 300 AA.
ID QY4U4;
AC QY4U4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE HISTONE DEACETYLASE 2 ISOFORM C.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OX Panicoideae; Andropogoneae; Zea.
RN NCBI_TaxID=4577;
RP SEQUENCE FROM N.A.
RA Dangl M., Brosch G., Haas H., Loidl P., Lusser A.;
RT "Molecular characterization of type-2 histone deacetylases in higher
plants.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254073; AAF68625.1; -.
DR InterPro; IPR000822; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2.1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
KW DNA-binding; Zinc-finger;
SQ SEQUENCE 300 AA; 32457 MW; 62P6C4148B1E1FW CRC64;

Query Match 33.5%; Score 532; DB 10; Length 300;
Best Local Similarity 44.0%; Pred. No. 3.9e-28;
Matches 124; Conservative 44; Mismatches 92; Indels 22; Gaps 9;

QY 1 MEFWGVAVTPKNAKVTPEEDSLVHISOASLDCTVKSQSVLVTGGAKLVIGTILSD 60
Db 1 MEFWGLEVKPGSTVKGPGYGVFLHLSQAALGESKKS-DNALMYVKIDQKLAIGTILSD 59

QY 61 KFOISDLVDFKFEFELSHSGTKANVHFICYKSPNIEQDDFTSSDDDEVDPEAVPAPAPTA 120
Db 60 KNPHIQDFLDFKFEFELSHSGTKSVFTTCYKVPQPEDEMDLSDDEDEELNVP---A 116

QY 121 VTANGNAGAAVVKADTKPKAKPAEYKPAEKPESEDESEDESEED--DDSEKGMV 178
Db 117 VKENGKADKKQKQSKAVAAAPSKSSPDKSKDDDDSDDETDSDDEGLSP 176

QY 179 DEEDSDDEDEEDSDDEEE--TPKKPEPINKKRPNE-SYKSTPVSCKKAKPAAPASTPQ 236
Db 177 EGGDDSDDEEDTSDDEEDTTPKPEYGVKRAESSVLKTPLSDKKAK-VATPSS--Q 233

QY 237 KTEKKKGG--HTATPHAKKGGKSPVNAQSPKS-----PKSG 269
Db 234 KTGCKKGAHVATPHPAK--GKTIVNNDKSVKSPKSAKSG 273

RESULT 9
QY4T4 PRELIMINARY; PRT; 294 AA.
ID QY4T4;
AC QY4T4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE PUTATIVE HISTONE DEACETYLASE HD2C (AT5G03740/f17C15_160).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsia.
RN NCBI_TaxID=3702;
RP SEQUENCE FROM N.A.

```

```

Query Match      28.4%; Score 450.5; DB 10; Length 296;
Best Local Similarity 38.3%; Pred. No. le-22;
Matches 118; Conservative 39; Mismatches 94; Indels 57; Gaps 12;

QY      1 MEFQGVAVTPKNAVKTVPEDSLVHISOASLDCTVKSGES-----VLSVTYGGAKLV 53
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db      1 MEFQGVKVKQPLHLDOPGLDLRHLVISOAVL-----GESKNNTPEIQLXVTYVSGDKLL 54
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

QY      54 IGTLSQDKFOIPISDLVDFDKFELSHSGTKANVHFIQYKSPNIEDQDPTSDDEED----- 108
      ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

```

Query Match	28.4%;	Score 450.5;	DB 10;	length 296;
Best Local Similarity	38.3%;	Pred. No. 1e-22;		
Matches 118:	Conservative	39;	Mismatches	94; Indels
				57; Gaps
QY	1	MEFWGVAATPKNATKVTPEEDSLVHISQASLDCTVKSGES-----VVLSVTYGGAKLV	53	
		: : : :		
Db	1	MEFWGVKKNGCKPHLDPGLRLVHISOVAL-----GESKNNVTEIDLYTVTVGSOKLL	54	
QY	54	IGTILSQDKFQIQLSDVLDFDKFELSHSGTKANHFICYKSPNIEQQDDFTSSDDDE----	108	
		: : : : : : : : : : :		
Db	55	IGTILSHEKFQQLSTEIVLERNFALSHTWKNGSVFFSGYKPEDLIDDQLEAGFKGWGLL	114	
QY	109	VPEAVPAPATVATANGACAAVVAADTKPAKPAEVKPAEKPESEDESDEDEESEE	168	
		: : : : : : : : : :		
Db	115	YPAAPKSAARQNVFQ-----LPNEVDKAK-----QDDADAGSEDSDDDDSENS	160	
QY	169	DDOSEKGM---DVDEDSDDEDEESEDSEETPKKPPEINNKRPNESVSKTPVSgkK	224	
		: :		
Db	161	GDEEEKVTAESEEDSDSDEEDDS---SEETPKKPPEEPKRSAEPNSKNPNASKK	217	
QY	225	AKPAAPASTPQKTEKKKG-G-HTATPHPAKGGKSPYNANOSPKSGSGSGGNNNK--P	28	
		: : : : : : : : : :		
Db	218	AK-----FVIPQKTDSKKPHVHVATPHPSQAQGN-----SGSGSTGEISKQOQT	26	
QY	282	FNSCKQFG	289	
		: :		
Db	264	KSAG-AFG	270	
RESULT	11			
Q94F81		PRELIMINARY;	PRT;	286 AA.
ID	Q94F81			
AC	Q94F81;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	H02 TYPE HISTONE DEACETYLASE HDA106.			
GN	HDA106.			
OS	Zea mays (Maize).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; PACC clade			
OX	Panicoidae; Andropogoneae; Zea.			
RX	NCBI_TaxID=4577;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. B73;			
RA	Chandler V.L., Kaepler S.M., Kaepler H.F., Cone K.C.;			
RT	*Sequences from the Plant Chromatin Consortium (NSF plant Genome			
RT	project 9975930).*			

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=CV. B73;
RA Bergstrom D., Springer N.M., Schmitt L.T., Guthrie E., Sidorenko L.,
RA Selinger D., Kaeppler S.M., Cone K.C.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF384033; AAK67143.1;
SQ SEQUENCE 286 AA; 30808 MW; E295E5CDFE4DEF1A CRC64;

Query Match 22.9%; Score 364; DB 10; Length 286;
Best Local Similarity 33.1%; Pred. No. 5.6e-17;
Matches 106; Conservative 43; Mismatches 81; Indels 90; Gaps 13;

QY 1 MEFVGVAVTPKNAKVTPEEDSLVHLSQASLDCTVKSAGESVLSVTGAKLVIGTSLSD 60
DB 1 MEVGQGVKPGATVSCVKGGLVHLSQALGESKASENAILSDDKLVIGTSLVSE 60
QY 61 KFPQISDLVDFKPELHSGTKANVHFICYKSP-----NIPQDFTSSDDSDVPEAVP 114
DB 61 KHPQISDLVDFKPELHSGTKANVHFICYKSPVPLFESNSGCD---SSDEE----- 110
QY 115 APAPTAVTANGAAGVAVKADTKPKAKPAEVK-PAREKPESEDEE-----S 160
DB 111 -----LKTQIP-LQNEIKISAAPKAPAKDDDDVFIILAMMMHIIYSS 152
QY 161 DDEDE-----SEDDSEKGMVDDESDDEE-----EDSEDE---EEEE 198
DB 153 DDDDDFTTSDNSNEMSEEDSDDEMESEDDSDDEMGAGSDSDSDSGSEHTS 212
QY 199 TPKKPE-PINKKRPNEVSXKTPVSGKKAPAAAPASTPQTEKKKGHTATPHPAK---- 253
DB 213 APKKTDDVVGKKR---AKAAEPYKKAK-----SEQSSQKTGDKASTSHPAKQSIK 261
QY 254 -----KGGKSPVNAQSPKSG 269
DB 262 TPADKSKTKPTADKSPKSG 281

RESULT 12
Q9M4T3 PRELIMINARY; PRT; 203 AA.
ID Q9M4T3
AC Q9M4T3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PUTATIVE HISTONE DEACETYLASE HD2D.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RA Bangl M., Brosch G., Haas H., Loidl P., Graessle S., Lusser A.;
RT "Molecular characterization of type-2 histone deacetylases in higher plants."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF255713; AAF70198.1;
SQ SEQUENCE 203 AA; 22651 MW; 69A1DA5941D7A293 CRC64;

Query Match 16.5%; Score 262.5; DB 10; Length 203;
Best Local Similarity 30.9%; Pred. No. 2.2e-10;
Matches 80; Conservative 33; Mismatches 75; Indels 71; Gaps 11;

QY 1 MEFVGVAVTPKNAKVTPEEDSLVHLSQASLDCT--VKSAGESVLSVTGCGAK--LVIGT 56
DB 1 MEFVGVAVTPKNAKVTPEEDSLVHLSQASLDCT--VKSAGESVLSVTGCGAK--LVIGT 56
QY 57 LSQDKFPQISDLVDFKPELHSGTKANVHFICYKSPNI--RQDFTTSSDDSDVPEAVP 114
DB 57 LSQDKFPQISDLVDFKPELHSGTKANVHFICYKSPNI--RQDFTTSSDDSDVPEAVP 114

DB 61 LSQ-KFPQFSDLYLGHEFEISHNST--SSVYLIGVYTFDAFDELDEIDSDSE----- 111
QY 115 APAPTAVTANGAAGVAVKADTKPKAKPAEVKPAEKPESEDESEDESEDESEDESEK 174
DB 112 -----LDEYMEQQAALPQNEINPEDESDS-----EM 141
QY 175 GMDVDEDDSDDEEEDSEDEETPKPEIPKPKPNES--VSKTPVSGKKAPAAAPA 232
DB 142 GLDEDDSDS-----EEDVEAEAPLVKVPSPKMPNGAFETAK---GGKK----- 183
QY 233 STPQTEKKKKGGHTATPHP 251
DB 184 -----NKSSGKKRCPFP 196

RESULT 13
Q26711 PRELIMINARY; PRT; 312 AA.
ID Q26711
AC Q26711;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NOP44/46.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=ETRO 164;
RX MEDLINE=96279091; PubMed=8663171;
RA Das A., Peterson G.C., Kanner S.B., Frevert U., Parsons M.;
RT "A major tyrosine-phosphorylated protein of Trypanosoma brucei is a
RT nucleolar RNA-binding protein."
RL J. Biol. Chem. 271:15675-15681(1996).
DR EMBL: U53863; AAC47134.1;
SQ SEQUENCE 312 AA; 33222 MW; 2A88ED8D4416ER00 CHC64;

Query Match 16.2%; Score 258; DB 5; Length 312;
Best Local Similarity 24.5%; Pred. No. 7e-10;
Matches 83; Conservative 55; Mismatches 117; Indels 84; Gaps 12;

QY 3 FVGAVTPKNAKVTPEEDSLVHLSQASLDCTVKSAGESVLSVTGCGAKLVIGTISQDK- 61
DB 4 FYGVEVSAGQKVKPKIPEDHVLRLQIAVPANA-SG-AITLVISFOGKEFTIATLDPKRS 61
QY 62 FQISDFDLVDFKPELHSGT-KANVHFICYKSPNIEQDDFTSSDDSDVPEA-----V 113
DB 62 LPOMGIDLVLTAEQGTTLSATSGSVHLTGYYVQPVDDGSDFDDEDDLDEVDAMFMEYCV 121
QY 114 P-----APAPTAVTANGAAGVAVKADTKPKAKPAEVKPAEKPESEDESEDESE 166
DB 122 PPKRGKSAKFPVSEENSEANGDVGDDSDDEEG-IEKDNIGIDAGDEDDDDDEA 180
QY 167 EDDDESEKGMVDDEDDSDDEEEDSEDEETPKPKPEIPKPKPNESVSKTPVSGKAK 226
DB 181 YDEDDSD-----DDDDDDDDDDDDDDDDDDDE----- 207
QY 227 PAAAPASTPQTEKK-----KGGHTATPHPAKKGKSPVNAQSPKSG 269
DB 208 -----PASKVQRTFQOQFPGNNGRGPFGDNRNGNGNGRPGFGDNRG-----NKG 257
QY 270 GQSSGNKPKFNSGKQFGGSNN-----KGSNKGKKG 303
DB 258 GDSGNGNNGRNGRGGDRGNGRNGRNGRGGDRGNGRNGRNG 296

RESULT 14
O45181 PRELIMINARY; PRT; 798 AA.
ID O45181
AC O45181;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

